EXPEDITED PROCEDURE
EXAMINING GROUP 1652

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant(s):	Yu and Turner, Jr.	Group Art Unit:	1652
Application No.:	10/044,807	Examiner:	S. Swope
Filed:	1/11/2002	Atty. Docket No.:	LEX-0298-USA

Title: Human Protease Polynucleotides and Compositions Comprising the Same (As Amended)

#11
AG.J
7/23/03

RESPONSE TO OFFICE ACTION DATED FEBRUARY 5, 2003

Mail Stop AF
Assistant Commissioner for Patents
Alexandria, VA 22313

Sir:

Applicants acknowledge the receipt of the Office Action ("the Action") mailed on February 5, 2003 (Paper No. 8), which has been carefully reviewed and studied. Reexamination and reconsideration of the application is requested in view of the following remarks. In order to facilitate the Examiner's evaluation of the application, Applicants have attempted to address the rejections in Paper No. 8 in the same order in which they were originally raised.

A Petition for an Extension of Time of two months to and including July 5, 2003, which falls on a Saturday and is therefore extended until Monday, July 7, 2003 under 37 C.F.R. § 1.7, and authorization to deduct the fee as required under 37 C.F.R. § 1.17(a)(2) from Applicants' representatives Deposit Account are included. The response is thus timely filed. Applicants believe no fees in addition to the fee for the extension of time are due in connection with this response. However, the Commissioner is authorized to charge any underpayment or credit any overpayment to Deposit Account No. 50-0892.

RESPONSE

I. Status of the Claims

No claims have been canceled. No claims have been amended. No new claims have been added.

Claims 1-4 are therefore presently pending in the case. For the convenience of the Examiner, a clean copy of the pending claims is attached hereto as **Exhibit A**.

II. Rejection of Claims 1-4 Under 35 U.S.C. § 101

The Action first rejects claims 1-4 under 35 U.S.C. § 101, as allegedly lacking a patentable utility. Applicants respectfully traverse.

As set forth in Applicants' response mailed on November 12, 2002 ("the previous response") to the First Office Action in this case, which was mailed on August 12 2002 ("the First Action"), the present invention has a number of substantial and credible utilities, not the least of which is in forensic analysis, as described in the specification, at least at page 3, line 15, and from page 11, line 31 to page 12, line 27. As described in the specification at page 18, lines 3-27, the present sequences define a number of coding single nucleotide polymorphisms - specifically: a C/G polymorphism at position 2361 of SEQ ID NO:1, which can result in an aspartate or glutamate at amino acid position 787 of SEQ ID NO:2; a C/A polymorphism at position 2467 of SEQ ID NO:1, which can result in a leucine or isoleucine at amino acid position 823 of SEQ ID NO:2; a C/A polymorphism at position 2613 of SEQ ID NO:1, both of which result in an isoleucine at corresponding aa position 871 of SEQ ID NO:2; a C/T polymorphism at position 3141 of SEQ ID NO:1, both of which result in a serine at amino acid position 1047 of SEQ ID NO:2; a G/T polymorphism at position 3225 of SEQ ID NO:1, which can result in a glutamine or histidine at amino acid position 1075 of SEQ ID NO:2; a C/T polymorphism at position 3226 of SEQ ID NO:1, which can result in an arginine or tryptophan at amino acid position 1076 of SEQ ID NO:2; and an A/G polymorphism at position 4226 of SEQ ID NO:1, which can result in an aspartate or glycine at amino acid position 1409 of SEQ ID NO:2. As such polymorphisms, and particularly combinations of polymorphisms, are the basis for forensic analysis, which does not require any information at all about the ultimate biological function of the encoded protein, and is undoubtedly

a "real world" utility, the present sequences must in themselves be useful.

The Examiner questions this asserted utility, stating "the presence of polymorphisms in human DNA is well established and virtually any locus on a human chromosome will exhibit one or more polymorphisms which could be so used" (Action at page 2). However, it is important to note that the presence of other polymorphic markers for forensic analysis does not mean that the present sequences lack a specific utility. As clearly stated by the Federal Circuit in *Carl Zeiss Stiftung v. Renishaw PLC*, 20 USPQ2d 1101 (Fed. Cir. 1991):

> An invention need not be the best or only way to accomplish a certain result, and it need only be useful to some extent and in certain applications: "[T]he fact that an invention has only limited utility and is only operable in certain applications is not grounds for finding a lack of utility." *Envirotech Corp. v. Al George, Inc.*, 221 USPQ 473, 480 (Fed. Cir. 1984)

Just because other polymorphic sequences from the human genome have been described does not mean that the use of the presently described polymorphic markers for forensic analysis is not a specific utility. The requirement for a **specific** utility, which is the proper standard for utility under 35 U.S.C. § 101, should not be confused with the requirement for a **unique** utility, which is clearly an improper standard. If every invention were required to have a unique utility, the Patent and Trademark Office would no longer be issuing patents on batteries, automobile tires, golf balls, golf clubs, and treatments for a variety of human diseases, just to name a few particular examples, because examples of each of these have already been described and patented. However, only the briefest perusal of any issue of the Official Gazette provides numerous examples of patents being granted on each of the above compositions every week. Furthermore, if a composition needed to be unique to be patented, the entire class and subclass system would be an effort in futility, as the class and subclass system serves solely to group such common inventions, which would not be required if each invention needed to have a unique utility. Thus, the present sequence clearly meets the requirements of 35 U.S.C. § 101.

The Examiner further states that "Applicants have not identified any particular reason for use of this particular polymorphism in forensic analysis or any particular benefit that would derive from analysis of this polymorphism" (Action at page 2). Applicants respectfully point out that the presently described polymorphisms are useful in forensic analysis for the same reason that any marker is useful in forensic analysis - specifically, to specifically identify individual members of the human population

based on the presence or absence of the described polymorphism. Using the polymorphic markers as described in the specification as originally field can distinguish members of a population from one another. In the worst case scenario, each of these markers are useful to distinguish 50% of the population (in other words, the marker being present in half of the population). The ability to eliminate 50% of the population from a forensic analysis **clearly** is a real world, practical utility. As set forth in *In re Langer* (183 USPQ 288 (CCPA 1974); "*Langer*"):

> As a matter of Patent Office practice, a specification which contains a disclosure of utility which corresponds in scope to the subject matter sought to be patented must be taken as sufficient to satisfy the utility requirement of § 101 for the entire claimed subject matter unless there is a reason for one skilled in the art to question the objective truth of the statement of utility or its scope.

Langer at 297, emphasis in original. As set forth in the MPEP, "Office personnel must provide evidence sufficient to show that the statement of asserted utility would be considered 'false' by a person of ordinary skill in the art" (MPEP, Eighth Edition at 2100-40, emphasis added). Thus, the present claims clearly meet the requirements of 35 U.S.C. § 101.

Furthermore, as the Examiner admits that the presently described polymorphism is a part of the family of polymorphisms that have a "well established" utility, the Federal Circuit's holding in *In re Brana,* (34 USPQ2d 1436 (Fed. Cir. 1995), "*Brana*") is directly on point. In *Brana*, the Federal Circuit admonished the Patent and Trademark Office for confusing "the requirements under the law for obtaining a patent with the requirements for obtaining government approval to market a particular drug for human consumption". *Brana* at 1442. The Federal Circuit went on to state:

> At issue in this case is an important question of the legal constraints on patent office examination practice and policy. The question is, with regard to pharmaceutical inventions, what must the applicant provide regarding the practical utility or usefulness of the invention for which patent protection is sought. This is not a new issue; it is one which we would have thought had been settled by case law years ago.

Brana at 1439, emphasis added. The choice of the phrase "utility or usefulness" in the foregoing quotation is highly pertinent. The Federal Circuit is evidently using "utility" to refer to rejections under 35 U.S.C. § 101, and is using "usefulness" to refer to rejections under 35 U.S.C. § 112, first paragraph. This is made evident in the continuing text in *Brana*, which explains the correlation between 35 U.S.C. §§ 101 and 112, first paragraph. The Federal Circuit concluded:

> FDA approval, however, is not a prerequisite for finding a compound useful within the meaning of the patent laws. Usefulness in patent law, and in particular in the context of pharmaceutical inventions, necessarily includes the expectation of further research and development. The stage at which an invention in this field becomes useful is well before it is ready to be administered to humans. Were we to require Phase II testing in order to prove utility, the associated costs would prevent many companies from obtaining patent protection on promising new inventions, thereby eliminating an incentive to pursue, through research and development, potential cures in many crucial areas such as the treatment of cancer.

Brana at 1442-1443, citations omitted, emphasis added. As set forth above, the present polymorphisms are useful in forensic analysis exactly as they are described in the specification as originally filed, without the need for any further research. Even if the use of these polymorphic markers provided additional information on the percentage of particular subpopulations that contain this polymorphic marker, this would not mean that "additional research" is needed in order for this marker as it is presently described in the instant specification to be of use to forensic science. As stated above, using the polymorphic marker as described in the specification as originally field can definitely distinguish members of a population from one another. However, even if, *arguendo*, further research might be required in certain aspects of the present invention, this does not preclude a finding that the invention has utility, as set forth by the Federal Circuit's holding in *Brana*, which clearly states, as highlighted in the quote above, that "pharmaceutical inventions, necessarily includes the expectation of further research and development" (*Brana* at 1442-1443, emphasis added). In assessing the question of whether undue experimentation would be required in order to practice the claimed invention, the key term is "undue", not "experimentation". *In re Angstadt and Griffin*, 190 USPQ 214 (CCPA 1976). The need for some experimentation does not render the claimed invention unpatentable. Indeed, a considerable amount of experimentation may be permissible if such experimentation is routinely practiced in the art. *In re Angstadt and Griffin, supra*; *Amgen, Inc. v. Chugai Pharmaceutical Co., Ltd.*, 18 USPQ2d 1016 (Fed. Cir. 1991). As a matter of law, it is well settled that a patent need not disclose what is well known in the art. *In re Wands*, 8 USPQ 2d 1400 (Fed. Cir. 1988).

Although Applicants need only make one credible assertion of utility to meet the requirements of 35 U.S.C. § 101 (*Raytheon v. Roper*, 220 USPQ 592 (Fed. Cir. 1983); *In re Gottlieb*, 140 USPQ 665 (CCPA 1964); *In re Malachowski*, 189 USPQ 432 (CCPA 1976); *Hoffman v. Klaus*, 9 USPQ2d 1657 (Bd. Pat. App. & Inter. 1988)), as set forth in the previous response, the present

sequence has a number of additional patentable utilities, among them, as detailed in the specification as originally filed, on page 3, lines 7-10, in "the identification of protein coding sequence". This is evidenced by the fact that SEQ ID NO:1 can be used to map the 29 coding exons on chromosome 9 (present within GenBank Accession Numbers AL591423, AL353895, AL449963, and AL158150, which are four overlapping clones from human chromosome 9; alignments and the first page from the GenBank records are shown in **Exhibit B**). The specification details, at page 3, lines 10-13, that the present sequence "identify biologically verified exon splice junctions, as opposed to splice junctions that may have been bioinformatically predicted from genomic sequence alone". It is well known that intron/exon boundaries are mutational hot spots, and thus the identification of the actual splice sites is of great utility to the skilled artisan. The specification details, at page 12, lines 5-11, that "sequences derived from regions adjacent to the intron/exon boundaries of the human gene can be used to design primers for use in amplification assays to detect mutations within the exons, introns, splice sites (*e.g.*, splice acceptor and/or donor sites), *etc.*, that can be used in diagnostics and pharmacogenomics". Applicants respectfully submit that the practical scientific value of biologically validated, expressed, spliced, and polyadenylated mRNA sequences is readily apparent to those skilled in the relevant biological and biochemical arts. Thus, the present claims clearly meet the requirements of 35 U.S.C. § 101.

As yet a further example of the utility of the presently claimed polynucleotides, as described in the specification at least at page 3, lines 7-8, the present nucleotide sequence has a specific utility in mapping the protein encoding regions of the corresponding human chromosome, specifically chromosome 9, as described in the specification at least on page 3, lines 8-10. This is evidenced by the fact that SEQ ID NO:1 can be used to map the 29 coding exons on chromosome 9, as detailed above (**Exhibit B**). Clearly, the present polynucleotide provides exquisite specificity in localizing the specific region of human chromosome 9 that contains the gene encoding the given polynucleotide, a utility not shared by virtually any other nucleic acid sequences. In fact, it is this specificity that makes this particular sequence so useful. Early gene mapping techniques relied on methods such as Giemsa staining to identify regions of chromosomes. However, such techniques produced genetic maps with a resolution of only 5 to 10 megabases, far too low to be of much help in identifying specific genes involved in disease. The skilled artisan readily appreciates the significant benefit afforded by markers

that map a specific locus of the human genome, such as the present nucleic acid sequence. For further evidence in support of the Applicants' position, the Examiner is invited to review, for example, section 3 of Venter *et al.* (2001, Science *291*:1304, at pp. 1317-1321, including Fig. 11 at pp.1324-1325), which demonstrates the significance of expressed sequence information in the structural analysis of genomic data. The presently claimed polynucleotide sequence defines a biologically validated sequence that provides a unique and specific resource for mapping the genome essentially as described in the Venter *et al.* article.

Applicants respectfully remind the Examiner that only a minor percentage (2-4%) of the genome actually encodes exons, which in-turn encode amino acid sequences. The presently claimed polynucleotide sequence provides biologically validated empirical data (*e.g.*, showing which sequences are transcribed, spliced, and polyadenylated) that *specifically* define that portion of the corresponding genomic locus that actually encodes exon sequence, as described above. Equally significant is that the claimed polynucleotide sequence defines how the encoded exons are actually spliced together to produce an active transcript (*i.e.*, the described sequences are useful for functionally defining exon splice-junctions). Thus, the present claims clearly meet the requirements of 35 U.S.C. § 101.

The Action also questions these asserted utilities, stating that "applicants have not identified any particular reason for using this polynucleotide in mapping chromosome 9" (Action bridging pages 3 and 4). The Examiner once again seems to be confusing the requirements of a **specific** utility with a **unique** utility. The fact that a small number of other nucleotide sequences could be used to map the protein coding regions in this specific region of chromosome 9 does not mean that the use of Applicants' sequence to map the protein coding regions of chromosome 9 is not a specific utility (*Carl Zeiss Stiftung v. Renishaw PLC, supra*).

In the previous response, Applicants detailed an additional example of the utility of the present nucleotide sequences, as described in the specification on page 6, lines 16-18, specifically that the present nucleotide sequences have utility in assessing gene expression patterns using high-throughput DNA chips. As previously set forth, evidence of the "real world" substantial utility of the present invention is further provided by the fact that there is an entire industry established based on the use of gene sequences or fragments from genes in a gene chip format. Perhaps the most notable gene chip company is Affymetrix. Affymetrix is clearly a "real world" company, as evidenced the fact that the

United States Patent and Trademark Office has issued numerous U.S. Patents to Affymetrix covering gene chip technology, as exemplified by U.S. Patent Nos. 5,445,934, 5,556,752, 5,744,305, 5,837,832, 6,156,501 and 6,261,776. However, there are many companies which have, at one time or another, concentrated on the use of gene sequences or fragments, in gene chip and non-gene chip formats, for example: Gene Logic, ABI-Perkin-Elmer, HySeq and Incyte. In addition, one such company (Rosetta Inpharmatics) was viewed to have such "real world" value that it was acquired by large a pharmaceutical company (Merck) for significant sums of money (net equity value of the transaction was $620 million). Given the widespread utility of such "gene chip" methods using non-biologically validated, *public domain* gene sequence information, there can be little doubt that the use of the presently described *novel* biologically validated coding sequence would have great utility in such DNA chip applications. The "real world" substantial industrial utility of gene sequences or fragments would, therefore, appear to be widespread and well established. Furthermore, compositions that enhance the utility of such DNA chips must in themselves be useful. Thus, the present claims clearly meet the requirements of 35 U.S.C. § 101.

The Action also questions this utility, stating that "Applicants have also not identified any particular reason for use of this particular polynucleotide in "DNA chips" (Action at page 2). First, Applicants point out that nucleic acid sequences are commonly used in gene chip applications without any information regarding the function of the encoded protein, or even evidence regarding whether the sequence is actually even expressed. Thus, the present sequence, which has been biologically validated to be expressed, has a much greater utility than sequences that are merely predicted to be expressed based on bioinformatic analysis. Additionally, Applicants point out that nucleic acid sequences such as SEQ ID NO:1 are routinely used by companies throughout the biotechnology sector **exactly** as they are presented in the Sequence Listing, without **any** further experimentation. Expression profiling does **not** require a knowledge of the function of the particular nucleic acid on the chip - rather the gene chip indicates which DNA fragments are expressed at greater or lesser levels in two or more particular tissue types. Furthermore, although further information regarding the biological activity of a particular nucleic acid sequence might make it even more useful in gene chip applications, this does not mean that the use of the presently claimed nucleic acid sequence in gene chip applications is not a specific utility (*Carl Zeiss Stiftung v. Renishaw PLC, supra*).

Additionally, Applicants point out that two sequences sharing nearly 100% percent identity at the protein level over an extended region of the claimed sequence is present in the leading scientific repository for biological sequence data (GenBank), and has been annotated by third party scientists *wholly unaffiliated with Applicants* as "Homo sapiens ADAMTS-like 1" variants 1 and 2 (GenBank accession numbers NM_139238 and NM_052866; alignments and GenBank reports are shown in **Exhibit C**). In the specification as originally filed, Applicants noted the similarity of the present sequence to "matrix metalloprotease" (specification at page 2, lines 7-8), and particularly "the ADAMTS family of metalloproteases" (specification at page 17, lines 31-32). Furthermore, the scientists that described ADAMTS-like 1 have determined that the protein is localized to the extracellular matrix (Hirohata *et al.*, J. Biol. Chem. *277*:12182-12189, 2002; **Exhibit D**). Applicants respectfully point out that the legal test for utility simply involves an assessment of whether those skilled in the art would find any of the utilities described for the invention to be believable. Given these two GenBank annotations and the manuscript by Hirohata *et al.*, there can be no question that those skilled in the art would clearly believe that Applicants' sequence is an ADAMTS-like protease, and would thus readily understand the utility of the presently claimed sequence, as described above, particularly in gene chip applications. As this is the standard for meeting the utility requirement of 35 U.S.C. § 101, Applicants submit that the present claims must clearly meet the requirements of 35 U.S.C. § 101.

Finally, as set forth in the previous response, the requirements set forth in the Action for compliance with 35 U.S.C. § 101 do not comply with the requirements set forth by the Patent and Trademark Office ("the PTO") itself for compliance with 35 U.S.C. § 101. While Applicants are well aware of the new Utility Guidelines set forth by the USPTO, Applicants respectfully point out that the current rules and regulations regarding the examination of patent applications is and always has been the patent laws as set forth in 35 U.S.C. and the patent rules as set forth in 37 C.F.R., not the Manual of Patent Examination Procedure or particular guidelines for patent examination set forth by the USPTO. Furthermore, it is the job of the judiciary, not the USPTO, to interpret these laws and rules. Applicants are unaware of any significant recent changes in either 35 U.S.C. § 101, or in the interpretation of 35 U.S.C. § 101 by the Supreme Court or the Federal Circuit that is in keeping with the new Utility Guidelines set forth by the USPTO. This is underscored by numerous patents that have been issued over the years that claim nucleic acid fragments that do not comply with the new Utility

Guidelines. As examples of such issued U.S. Patents, the Examiner is invited to review U.S. Patent Nos. 5,817,479, 5,654,173, and 5,552,281 (each of which claims short polynucleotides), and recently issued U.S. Patent No. 6,340,583 (which includes no working examples), none of which contain examples of the "real-world" utilities that the Examiner seems to be requiring. As issued U.S. Patents are presumed to meet all of the requirements for patentability, including 35 U.S.C. §§ 101 and 112, first paragraph (see Section III, below), Applicants submit that the present polynucleotides must also meet the requirements of 35 U.S.C. § 101. While Applicants understand that each application is examined on its own merits, Applicants are unaware of any changes to 35 U.S.C. § 101, or in the interpretation of 35 U.S.C. § 101 by the Supreme Court or the Federal Circuit, since the issuance of these patents that render the subject matter claimed in these patents, which is similar to the subject matter in question in the present application, as suddenly non-statutory or failing to meet the requirements of 35 U.S.C. § 101. Thus, holding Applicants to a different standard of utility would be arbitrary and capricious, and, like other clear violations of due process, cannot stand.

For each of the foregoing reasons, as well as the reasons set forth in the previous response, Applicants submit that as the presently claimed nucleic acid molecules have been shown to have a substantial, specific, credible and well-established utility, the rejection of claims 1-4 under 35 U.S.C. § 101 has been overcome, and request that the rejection be withdrawn.

III. Rejection of Claims 1-4 Under 35 U.S.C. § 112, First Paragraph

The Action next rejects claims 1-4 under 35 U.S.C. § 112, first paragraph, since allegedly one skilled in the art would not know how to use the invention, as the invention allegedly is not supported by a specific, substantial, and credible utility or a well-established utility. Applicants respectfully traverse.

Applicants submit that as claims 1-4 have been shown to have "a specific, substantial, and credible utility", as detailed in section II above, the present rejection of claims 1-4 under 35 U.S.C. § 112, first paragraph, cannot stand.

Applicants therefore request that the rejection of claims 1-4under 35 U.S.C. § 112, first paragraph, be withdrawn.

IV. Conclusion

The present document is a full and complete response to the Action. In conclusion, Applicants submit that, in light of the foregoing remarks, the present case is in condition for allowance, and such favorable action is respectfully requested. Should Examiner Swope have any questions or comments, or believe that certain amendments of the claims might serve to improve their clarity, a telephone call to the undersigned Applicants' representative is earnestly solicited.

Respectfully submitted,

July 7, 2003
Date

David W. Hibler Reg. No. 41,071
Agent for Applicants

LEXICON GENETICS INCORPORATED
8800 Technology Forest Place
The Woodlands, TX 77381
(281) 863-3399

Query= hMEM_224_ORF
(5289 letters)

>AL591423.6.1.54193
Length = 54193

Score = 2218 bits (1119), Expect = 0.0
Identities = 1125/1127 (99%)
Strand = Plus / Plus

```
Query: 2552  ggcccgggcggccatccacgaagcacagcccgcacatcgcggccgccaggaaggtctaca 2611
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 25657 ggcccgggcggccatccacgaagcacagcccgcacatcgcggccgccaggaaggtctaca 25716

Query: 2612  tccagactcgcaggcagaggaagctgcacttcgtggtgggggcttcgcctacctgctcc 2671
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 25717 tacagactcgcaggcagaggaagctgcacttcgtggtgggggcttcgcctacctgctcc 25776

Query: 2672  ccaagacggcggtggtgctgcgctgcccggcgcgcagggtccgcaagcccctcatcacct 2731
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 25777 ccaagacggcggtggtgctgcgctgcccggcgcgcagggtccgcaagcccctcatcacct 25836

Query: 2732  gggagaaggacggccagcacctcatcagctcgacgcacgtcacggtggccccttcggct 2791
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 25837 gggagaaggacggccagcacctcatcagctcgacgcacgtcacggtggccccttcggct 25896

Query: 2792  atctcaagatccaccgcctcaagccctcggatgcaggcgtctacacctgctcagcgggcc 2851
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 25897 atctcaagatccaccgcctcaagccctcggatgcaggcgtctacacctgctcagcgggcc 25956

Query: 2852  cggcccgggagcactttgtgattaagctcatcggaggcaaccgcaagctcgtggcccggc 2911
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 25957 cggcccgggagcactttgtgattaagctcatcggaggcaaccgcaagctcgtggcccggc 26016

Query: 2912  ccttgagcccgagaagtgaggaagaggtgcttgcggggaggaagggcggcccgaaggagg 2971
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 26017 ccttgagcccgagaagtgaggaagaggtgcttgcggggaggaagggcggcccgaaggagg 26076

Query: 2972  ccctgcagacccacaaacaccagaacgggatcttctccaacggcagcaaggcggagaagc 3031
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 26077 ccctgcagacccacaaacaccagaacgggatcttctccaacggcagcaaggcggagaagc 26136

Query: 3032  ggggcctggccgccaacccggggagccgctacgacgacctcgtctcccggctgctggagc 3091
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 26137 ggggcctggccgccaacccggggagccgctacgacgacctcgtctcccggctgctggagc 26196
```

```
Query: 3092  agggcggctggcccggagagctgctggcctcgtgggaggcgcaggactccgcggaaagga 3151
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Sbjct: 26197 agggcggctggcccggagagctgctggcctcgtgggaggcgcaggactctgcggaaagga 26256

Query: 3152  acacgacctcggaggaggacccgggtgcagagcaagtgctcctgcacctgcccttcacca 3211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 26257 acacgacctcggaggaggacccgggtgcagagcaagtgctcctgcacctgcccttcacca 26316

Query: 3212  tggtgaccgagcagcggcgcctggacgacatcctggggaacctctcccagcagcccgagg 3271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 26317 tggtgaccgagcagcggcgcctggacgacatcctggggaacctctcccagcagcccgagg 26376

Query: 3272  agctgcgcgacctctacagcaagcacctggtggcccagctggcccaggagatcttccgca 3331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 26377 agctgcgcgacctctacagcaagcacctggtggcccagctggcccaggagatcttccgca 26436

Query: 3332  gccacctggagcaccaggacacgctcctgaagccctcggagcgcaggacttccccagtga 3391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 26437 gccacctggagcaccaggacacgctcctgaagccctcggagcgcaggacttccccagtga 26496

Query: 3392  ctctctcgcctcataaacacgtgtctggcttcagcagctccctgcggacctcctccaccg 3451
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 26497 ctctctcgcctcataaacacgtgtctggcttcagcagctccctgcggacctcctccaccg 26556

Query: 3452  gggacgccgggggaggctctcgaaggccacaccgcaagcccaccatcctgcgcaagatct 3511
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 26557 gggacgccgggggaggctctcgaaggccacaccgcaagcccaccatcctgcgcaagatct 26616

Query: 3512  cagcggcccagcagctctcagcctcggaggtggtcacccacctggggcagacggtggccc 3571
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 26617 cagcggcccagcagctctcagcctcggaggtggtcacccacctggggcagacggtggccc 26676

Query: 3572  tggccagcgggacactgagtgttcttctgcactgtgaggccatcggccacccaaggccta 3631
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 26677 tggccagcgggacactgagtgttcttctgcactgtgaggccatcggccacccaaggccta 26736

Query: 3632  ccatcagctgggccaggaatggagaagaagttcagttcagtgacagg 3678
              ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 26737 ccatcagctgggccaggaatggagaagaagttcagttcagtgacagg 26783
```

```
 Score =  422 bits (213), Expect = e-115
 Identities = 213/213 (100%)
 Strand = Plus / Plus


Query: 2005 aggtgggaaattggcaagtggagtccatgtagtctcacatgtggggtcggcctacagacc 2064
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 2172 aggtgggaaattggcaagtggagtccatgtagtctcacatgtggggtcggcctacagacc 2231


Query: 2065 agagacgtcttctgcagccacctgctttccagagagatgaatgaaacagtcatcctggct 2124
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 2232 agagacgtcttctgcagccacctgctttccagagagatgaatgaaacagtcatcctggct 2291


Query: 2125 gatgagctgtgtcgccagcccaagcccagcacggtgcaagcttgtaaccgctttaattgc 2184
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 2292 gatgagctgtgtcgccagcccaagcccagcacggtgcaagcttgtaaccgctttaattgc 2351


Query: 2185 cccccagcctggtaccctgcacagtggcagccg 2217
           |||||||||||||||||||||||||||||||
Sbjct: 2352 cccccagcctggtaccctgcacagtggcagccg 2384



 Score =  359 bits (181), Expect = 1e-95
 Identities = 181/181 (100%)
 Strand = Plus / Plus


Query: 2217  gtgttccagaacgtgtggcgggggtgttcagaaacgtgaggttctttgcaagcagcgcat 2276
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 19477 gtgttccagaacgtgtggcgggggtgttcagaaacgtgaggttctttgcaagcagcgcat 19536


Query: 2277  ggctgatggcagcttcctggagcttcctgagaccttctgttcagcttcaaaacctgcctg 2336
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 19537 ggctgatggcagcttcctggagcttcctgagaccttctgttcagcttcaaaacctgcctg 19596


Query: 2337  ccagcaagcatgcaagaaagatgactgtcccagcgagtggcttctctcagactggacaga 2396
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 19597 ccagcaagcatgcaagaaagatgactgtcccagcgagtggcttctctcagactggacaga 19656


Query: 2397  g 2397
              |
Sbjct: 19657 g 19657
```

```
 Score =  311 bits (157), Expect = 2e-81
 Identities = 157/157 (100%)
 Strand = Plus / Plus


Query: 2396  agtgttccacaagctgcggggaaggcacccagactcgaagcgccatttgccgaaagatgc 2455
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 24617 agtgttccacaagctgcggggaaggcacccagactcgaagcgccatttgccgaaagatgc 24676


Query: 2456  tgaaaaccggcctctcaacggttgtcaattccaccctgtgcccgcccctgcctttctctt 2515
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 24677 tgaaaaccggcctctcaacggttgtcaattccaccctgtgcccgcccctgcctttctctt 24736


Query: 2516  cctccatcaggccctgtatgctggcaacctgtgcaag 2552
              |||||||||||||||||||||||||||||||||||
Sbjct: 24737 cctccatcaggccctgtatgctggcaacctgtgcaag 24773



 Score =  262 bits (132), Expect = 2e-66
 Identities = 132/132 (100%)
 Strand = Plus / Plus


Query: 3675  caggattcttctacagccagatgattccttacagatcttggcaccagtggaagcagatgt 3734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 44270 caggattcttctacagccagatgattccttacagatcttggcaccagtggaagcagatgt 44329


Query: 3735  gggtttctacacttgcaatgccaccaatgccttgggatacgactctgtctccattgccgt 3794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 44330 gggtttctacacttgcaatgccaccaatgccttgggatacgactctgtctccattgccgt 44389


Query: 3795  cacattagcagg 3806
              ||||||||||||
Sbjct: 44390 cacattagcagg 44401
```

```
>AL353895.4.1.163163
          Length = 163163

 Score =  603 bits (304), Expect = e-169
 Identities = 304/304 (100%)
 Strand = Plus / Plus


Query: 1575   gttcatcccagaggcctggtcggcctgcacagtcacctgtggtgtggggacccaggtgcg 1634
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 116786 gttcatcccagaggcctggtcggcctgcacagtcacctgtggtgtggggacccaggtgcg 116845


Query: 1635   aatagtcaggtgccaggtgctcctgtctttctctcagtccgtggctgacctgcctattga 1694
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 116846 aatagtcaggtgccaggtgctcctgtctttctctcagtccgtggctgacctgcctattga 116905


Query: 1695   cgagtgtgaagggcccaagccagcatcccagcgtgcctgttatgcaggcccatgcagcgg 1754
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 116906 cgagtgtgaagggcccaagccagcatcccagcgtgcctgttatgcaggcccatgcagcgg 116965


Query: 1755   ggaaattcctgagttcaacccagacgagacagatgggctctttggtggcctgcaggattt 1814
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 116966 ggaaattcctgagttcaacccagacgagacagatgggctctttggtggcctgcaggattt 117025


Query: 1815   cgacgagctgtatgactgggagtatgaggggttcaccaagtgctccgagtcctgtggagg 1874
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 117026 cgacgagctgtatgactgggagtatgaggggttcaccaagtgctccgagtcctgtggagg 117085


Query: 1875   aggt 1878
              ||||
Sbjct: 117086 aggt 117089



 Score =  408 bits (206), Expect = e-110
 Identities = 206/206 (100%)
 Strand = Plus / Plus


Query: 1136  ggtgggaggccaccccatggaccgcgtgctcctcctcgtgtgggggggggcatccagagcc 1195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 90348 ggtgggaggccaccccatggaccgcgtgctcctcctcgtgtgggggggggcatccagagcc 90407


Query: 1196  gggcagtttcctgtgtggaggaggacatccaggggcatgtcacttcagtggaagagtgga 1255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 90408 gggcagtttcctgtgtggaggaggacatccaggggcatgtcacttcagtggaagagtgga 90467


Query: 1256  aatgcatgtacacccctaagatgcccatcgcgcagccctgcaacatttttgactgcccta 1315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 90468 aatgcatgtacacccctaagatgcccatcgcgcagccctgcaacatttttgactgcccta 90527
```

```
Query: 1316  aatggctggcacaggagtggtctccg 1341
             |||||||||||||||||||||||||||
Sbjct: 90528 aatggctggcacaggagtggtctccg 90553
```

Score = 305 bits (154), Expect = 2e-79
Identities = 157/158 (99%)
Strand = Plus / Plus

```
Query: 677   atctggaaaccaaaaccctccaggggactaaaggtgaaaacagtctcagctccacaggaa 736
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Sbjct: 49286 atctggaaaccaaaaccctccaggggactaaaggtgaaaacagtctcaactccacaggaa 49345

Query: 737   ctttccttgtggacaattctagtgtggacttccagaaatttccagacaaagagatactga 796
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 49346 ctttccttgtggacaattctagtgtggacttccagaaatttccagacaaagagatactga 49405

Query: 797   gaatggctggaccactcacagcagatttcattgtcaag 834
             |||||||||||||||||||||||||||||||||||||
Sbjct: 49406 gaatggctggaccactcacagcagatttcattgtcaag 49443
```

Score = 295 bits (149), Expect = 1e-76
Identities = 149/149 (100%)
Strand = Plus / Plus

```
Query: 1341  gtgcacagtgacatgtggccagggcctcagataccgtgtggtcctctgcatcgaccatcg 1400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 91850 gtgcacagtgacatgtggccagggcctcagataccgtgtggtcctctgcatcgaccatcg 91909

Query: 1401  aggaatgcacacaggaggctgtagcccaaaaacaaagccccacataaaagaggaatgcat 1460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 91910 aggaatgcacacaggaggctgtagcccaaaaacaaagccccacataaaagaggaatgcat 91969

Query: 1461  cgtacccactccctgctataaacccaaag 1489
             |||||||||||||||||||||||||||
Sbjct: 91970 cgtacccactccctgctataaacccaaag 91998
```

Score = 280 bits (141), Expect = 9e-72
Identities = 141/141 (100%)
Strand = Plus / Plus

```
Query: 945   aggttatcagctgacatcggctgagtgctacgatctgaggagcaaccgtgtggttgctga 1004
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 71970 aggttatcagctgacatcggctgagtgctacgatctgaggagcaaccgtgtggttgctga 72029
```

```
Query: 1005  ccaatactgtcactattacccagagaacatcaaacccaaacccaagcttcaggagtgcaa 1064
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 72030 ccaatactgtcactattacccagagaacatcaaacccaaacccaagcttcaggagtgcaa 72089


Query: 1065  cttggatccttgtccagccag 1085
              ||||||||||||||||||||
Sbjct: 72090 cttggatccttgtccagccag 72110
```

Score = 266 bits (134), Expect = 1e-67
Identities = 134/134 (100%)
Strand = Plus / Plus

```
Query: 1875   aggtgtccaggaggctgtggtgagctgcttgaacaaacagactcgggagcctgctgagga 1934
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 131573 aggtgtccaggaggctgtggtgagctgcttgaacaaacagactcgggagcctgctgagga 131632


Query: 1935   gaacctgtgcgtgaccagccgccggcccccacagctcctgaagtcctgcaatttggatcc 1994
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 131633 gaacctgtgcgtgaccagccgccggcccccacagctcctgaagtcctgcaatttggatcc 131692


Query: 1995   ctgcccagcaaggt 2008
               ||||||||||||||
Sbjct: 131693 ctgcccagcaaggt 131706
```

Score = 252 bits (127), Expect = 2e-63
Identities = 127/127 (100%)
Strand = Plus / Plus

```
Query: 475   attgttggctgcgatcaccagctgggaagcaccgtcaaggaagataactgtggggtctgc 534
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 32280 attgttggctgcgatcaccagctgggaagcaccgtcaaggaagataactgtggggtctgc 32339


Query: 535   aacggagatgggtccacctgccggctggtccgagggcagtataaatcccagctctccgca 594
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 32340 aacggagatgggtccacctgccggctggtccgagggcagtataaatcccagctctccgca 32399


Query: 595   accaaat 601
              |||||||
Sbjct: 32400 accaaat 32406
```

```
 Score =  230 bits (116), Expect = 7e-57
 Identities = 116/116 (100%)
 Strand = Plus / Plus


Query: 833   agattcgtaactcgggctccgctgacagtacagtccagttcatcttctatcaacccatca 892
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 67675 agattcgtaactcgggctccgctgacagtacagtccagttcatcttctatcaacccatca 67734


Query: 893   tccaccgatggagggagacggatttctttccttgctcagcaacctgtggaggaggt 948
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 67735 tccaccgatggagggagacggatttctttccttgctcagcaacctgtggaggaggt 67790


 Score =  176 bits (89), Expect = 9e-41
 Identities = 89/89 (100%)
 Strand = Plus / Plus


Query: 1488  agagaaacttccagtcgaggccaagttgccatggttcaaacaagctcaagagctagaaga 1547
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 94753 agagaaacttccagtcgaggccaagttgccatggttcaaacaagctcaagagctagaaga 94812


Query: 1548  aggagctgctgtgtcagaggagccctcgt 1576
              ||||||||||||||||||||||||||||
Sbjct: 94813 aggagctgctgtgtcagaggagccctcgt 94841


 Score =  149 bits (75), Expect = 2e-32
 Identities = 75/75 (100%)
 Strand = Plus / Plus


Query: 602   cggatgatactgtggttgcaattccctatggaagtagacatattcgccttgtcttaaaag 661
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 45975 cggatgatactgtggttgcaattccctatggaagtagacatattcgccttgtcttaaaag 46034


Query: 662   gtcctgatcacttat 676
              |||||||||||||||
Sbjct: 46035 gtcctgatcacttat 46049


 Score =  111 bits (56), Expect = 5e-21
 Identities = 56/56 (100%)
 Strand = Plus / Plus


Query: 1083  cagtgacggatacaagcagatcatgccttatgacctctaccatccccttcctcggt 1138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 85891 cagtgacggatacaagcagatcatgccttatgacctctaccatccccttcctcggt 85946
```

>AL449963 ACCESSION:AL449963 NID: gi 20387012 emb AL449963.2 HS399M15 Homo sapiens chromosome 9 BAC RP11-399M15, complete sequence
Length = 213216

Score = 472 bits (238), Expect = e-129
Identities = 238/238 (100%)
Strand = Plus / Plus

```
Query: 237    ggactgcccaccagaagcaggtgatttccgagctcagcaatgctcagctcataatgatgt 296
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 104804 ggactgcccaccagaagcaggtgatttccgagctcagcaatgctcagctcataatgatgt 104863


Query: 297    caagcaccatggccagttttatgaatggcttcctgtgtctaatgaccctgacaacccatg 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 104864 caagcaccatggccagttttatgaatggcttcctgtgtctaatgaccctgacaacccatg 104923


Query: 357    ttcactcaagtgccaagccaaaggaacaaccctggttgttgaactagcacctaaggtctt 416
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 104924 ttcactcaagtgccaagccaaaggaacaaccctggttgttgaactagcacctaaggtctt 104983


Query: 417    agatggtacgcgttgctatacagaatctttggatatgtgcatcagtggtttatgccaa 474
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 104984 agatggtacgcgttgctatacagaatctttggatatgtgcatcagtggtttatgccaa 105041
```

Score = 408 bits (206), Expect = e-110
Identities = 206/206 (100%)
Strand = Plus / Plus

```
Query: 1136   ggtgggaggccaccccatggaccgcgtgctcctcctcgtgtgggggggcatccagagcc 1195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 211086 ggtgggaggccaccccatggaccgcgtgctcctcctcgtgtgggggggcatccagagcc 211145


Query: 1196   gggcagtttcctgtgtggaggaggacatccaggggcatgtcacttcagtggaagagtgga 1255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 211146 gggcagtttcctgtgtggaggaggacatccaggggcatgtcacttcagtggaagagtgga 211205


Query: 1256   aatgcatgtacacccctaagatgcccatcgcgcagccctgcaacatttttgactgcccta 1315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 211206 aatgcatgtacacccctaagatgcccatcgcgcagccctgcaacatttttgactgcccta 211265


Query: 1316   aatggctggcacaggagtggtctccg 1341
              |||||||||||||||||||||||||
Sbjct: 211266 aatggctggcacaggagtggtctccg 211291
```

```
 Score =  313 bits (158), Expect = 2e-81
 Identities = 158/158 (100%)
 Strand = Plus / Plus


Query: 677    atctggaaaccaaaaccctccaggggactaaaggtgaaaacagtctcagctccacaggaa 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 170029 atctggaaaccaaaaccctccaggggactaaaggtgaaaacagtctcagctccacaggaa 170088


Query: 737    ctttccttgtggacaattctagtgtggacttccagaaatttccagacaaagagatactga 796
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 170089 ctttccttgtggacaattctagtgtggacttccagaaatttccagacaaagagatactga 170148


Query: 797    gaatggctggaccactcacagcagatttcattgtcaag 834
              |||||||||||||||||||||||||||||||||||||
Sbjct: 170149 gaatggctggaccactcacagcagatttcattgtcaag 170186



 Score =  295 bits (149), Expect = 4e-76
 Identities = 149/149 (100%)
 Strand = Plus / Plus


Query: 1341   gtgcacagtgacatgtggccagggcctcagataccgtgtggtcctctgcatcgaccatcg 1400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 212586 gtgcacagtgacatgtggccagggcctcagataccgtgtggtcctctgcatcgaccatcg 212645


Query: 1401   aggaatgcacacaggaggctgtagcccaaaaacaaagccccacataaaagaggaatgcat 1460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 212646 aggaatgcacacaggaggctgtagcccaaaaacaaagccccacataaaagaggaatgcat 212705


Query: 1461   cgtacccactccctgctataaacccaaag 1489
              |||||||||||||||||||||||||||
Sbjct: 212706 cgtacccactccctgctataaacccaaag 212734



 Score =  280 bits (141), Expect = 2e-71
 Identities = 141/141 (100%)
 Strand = Plus / Plus


Query: 945    aggttatcagctgacatcggctgagtgctacgatctgaggagcaaccgtgtggttgctga 1004
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 192708 aggttatcagctgacatcggctgagtgctacgatctgaggagcaaccgtgtggttgctga 192767


Query: 1005   ccaatactgtcactattacccagagaacatcaaacccaaacccaagcttcaggagtgcaa 1064
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 192768 ccaatactgtcactattacccagagaacatcaaacccaaacccaagcttcaggagtgcaa 192827
```

```
Query: 1065   cttggatccttgtccagccag 1085
              ||||||||||||||||||||
Sbjct: 192828 cttggatccttgtccagccag 192848
```

Score = 258 bits (130), Expect = 8e-65
Identities = 130/130 (100%)
Strand = Plus / Plus

```
Query: 63    gagttccaggaccgcacgctccgaggaggaccgggacggcctatgggatgcctggggccc 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 35606 gagttccaggaccgcacgctccgaggaggaccgggacggcctatgggatgcctggggccc 35665

Query: 123   atggagtgaatgctcacgcacctgcgggggtggggcctcctactctctgaggcgctgcct 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 35666 atggagtgaatgctcacgcacctgcgggggtggggcctcctactctctgaggcgctgcct 35725

Query: 183   gagcagcaag 192
              ||||||||||
Sbjct: 35726 gagcagcaag 35735
```

Score = 252 bits (127), Expect = 5e-63
Identities = 127/127 (100%)
Strand = Plus / Plus

```
Query: 475    attgttggctgcgatcaccagctgggaagcaccgtcaaggaagataactgtggggtctgc 534
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 153018 attgttggctgcgatcaccagctgggaagcaccgtcaaggaagataactgtggggtctgc 153077

Query: 535    aacggagatgggtccacctgccggctggtccgagggcagtataaatcccagctctccgca 594
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 153078 aacggagatgggtccacctgccggctggtccgagggcagtataaatcccagctctccgca 153137

Query: 595    accaaat 601
              |||||||
Sbjct: 153138 accaaat 153144
```

Score = 230 bits (116), Expect = 2e-56
Identities = 116/116 (100%)
Strand = Plus / Plus

```
Query: 833    agattcgtaactcgggctccgctgacagtacagtccagttcatcttctatcaacccatca 892
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 188412 agattcgtaactcgggctccgctgacagtacagtccagttcatcttctatcaacccatca 188471
```

```
Query: 893    tccaccgatggagggagacggatttctttccttgctcagcaacctgtggaggaggt 948
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 188472 tccaccgatggagggagacggatttctttccttgctcagcaacctgtggaggaggt 188527


 Score =  149 bits (75), Expect = 5e-32
 Identities = 75/75 (100%)
 Strand = Plus / Plus


Query: 602    cggatgatactgtggttgcaattccctatggaagtagacatattcgccttgtcttaaaag 661
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 166718 cggatgatactgtggttgcaattccctatggaagtagacatattcgccttgtcttaaaag 166777


Query: 662    gtcctgatcacttat 676
              |||||||||||||||
Sbjct: 166778 gtcctgatcacttat 166792


 Score =  125 bits (63), Expect = 8e-25
 Identities = 63/63 (100%)
 Strand = Plus / Plus


Query: 1    atggaatgctgccgtcgggcaactcctggcacactgctcctctttctggctttcctgctc 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 5010 atggaatgctgccgtcgggcaactcctggcacactgctcctctttctggctttcctgctc 5069


Query: 61   ctg 63
             |||
Sbjct: 5070 ctg 5072


 Score =  111 bits (56), Expect = 1e-20
 Identities = 56/56 (100%)
 Strand = Plus / Plus


Query: 1083   cagtgacggatacaagcagatcatgccttatgacctctaccatccccttcctcggt 1138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 206629 cagtgacggatacaagcagatcatgccttatgacctctaccatccccttcctcggt 206684


 Score = 93.7 bits (47), Expect = 3e-15
 Identities = 47/47 (100%)
 Strand = Plus / Plus


Query: 192   gagctgtgaaggaagaaatatccgatacagaacatgcagtaatgtgg 238
              |||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 64022 gagctgtgaaggaagaaatatccgatacagaacatgcagtaatgtgg 64068
```

```
>AL158150.14.1.168011
          Length = 168011

 Score =  442 bits (223), Expect = e-120
 Identities = 223/223 (100%)
 Strand = Plus / Plus


Query: 4960   aggcctgtgagcacccagaactgctggtcagaggcctgcagtgtacactggagagtcagc 5019
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 103373 aggcctgtgagcacccagaactgctggtcagaggcctgcagtgtacactggagagtcagc 103432


Query: 5020   ctgtggaccctgtgcacagctacctgtggcaactacggcttccagtcccggcgtgtggag 5079
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 103433 ctgtggaccctgtgcacagctacctgtggcaactacggcttccagtcccggcgtgtggag 103492


Query: 5080   tgtgtgcatgcccgcaccaacaaggcagtgcctgagcacctgtgctcctgggggccccgg 5139
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 103493 tgtgtgcatgcccgcaccaacaaggcagtgcctgagcacctgtgctcctgggggccccgg 103552


Query: 5140   cctgccaactggcagcgctgcaacatcaccccatgtgaaaaca 5182
              |||||||||||||||||||||||||||||||||||||||||||
Sbjct: 103553 cctgccaactggcagcgctgcaacatcaccccatgtgaaaaca 103595



 Score =  424 bits (214), Expect = e-115
 Identities = 214/214 (100%)
 Strand = Plus / Plus


Query: 4249  ggctgccccatcaaaggtcaccctgtccctaatatcacctggtttcatggtggtcagcca 4308
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 84513 ggctgccccatcaaaggtcaccctgtccctaatatcacctggtttcatggtggtcagcca 84572


Query: 4309  attgtcactgccacaggactgacgcatcacatcttggcagctggacagatccttcaagtt 4368
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 84573 attgtcactgccacaggactgacgcatcacatcttggcagctggacagatccttcaagtt 84632


Query: 4369  gcaaaccttagcggtgggtctcaaggggaattcagctgccttgctcagaatgaggcaggg 4428
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 84633 gcaaaccttagcggtgggtctcaaggggaattcagctgccttgctcagaatgaggcaggg 84692


Query: 4429  gtgctcatgcagaaggcatctttagtgatccaag 4462
              |||||||||||||||||||||||||||||||||
Sbjct: 84693 gtgctcatgcagaaggcatctttagtgatccaag 84726
```

```
 Score =  414 bits (209), Expect = e-112
 Identities = 209/209 (100%)
 Strand = Plus / Plus


Query: 4643  ggtggatggtgacctcctggtctgcctgtacccggagctgtgggggaggtgtccagaccc 4702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 89071 ggtggatggtgacctcctggtctgcctgtacccggagctgtgggggaggtgtccagaccc 89130


Query: 4703  gcagggtgacctgtcaaaagctgaaagcctctgggatctccacccctgtgtccaatgaca 4762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 89131 gcagggtgacctgtcaaaagctgaaagcctctgggatctccacccctgtgtccaatgaca 89190


Query: 4763  tgtgcacccaggtcgccaagcggcctgtggacacccaggcctgtaaccagcagctgtgtg 4822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 89191 tgtgcacccaggtcgccaagcggcctgtggacacccaggcctgtaaccagcagctgtgtg 89250


Query: 4823  tggagtgggccttctccagctggggccag 4851
              |||||||||||||||||||||||||||||
Sbjct: 89251 tggagtgggccttctccagctggggccag 89279



 Score =  369 bits (186), Expect = 1e-98
 Identities = 186/186 (100%)
 Strand = Plus / Plus


Query: 4461  agattactggtggtctgtggacagactggcaacctgctcagcctcctgtggtaaccgggg 4520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 86249 agattactggtggtctgtggacagactggcaacctgctcagcctcctgtggtaaccgggg 86308


Query: 4521  ggttcagcagccccgcttgaggtgcctgctgaacagcacggaggtcaaccctgcccactg 4580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 86309 ggttcagcagccccgcttgaggtgcctgctgaacagcacggaggtcaaccctgcccactg 86368


Query: 4581  cgcagggaaggttcgccctgcggtgcagcccatcgcgtgcaaccggagagactgcccttc 4640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 86369 cgcagggaaggttcgccctgcggtgcagcccatcgcgtgcaaccggagagactgcccttc 86428


Query: 4641  tcggtg 4646
              ||||||
Sbjct: 86429 tcggtg 86434
```

```
 Score =  361 bits (182), Expect = 3e-96
 Identities = 182/182 (100%)
 Strand = Plus / Plus


Query: 3933  aggagtgcctgaagctgaagtcacttggttcaggaataaaagcaaactgggctccccgca 3992
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 22965 aggagtgcctgaagctgaagtcacttggttcaggaataaaagcaaactgggctccccgca 23024


Query: 3993  ccatctgcacgaaggctccttgctgctcacaaacgtgtcctcctcggatcagggcctgta 4052
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 23025 ccatctgcacgaaggctccttgctgctcacaaacgtgtcctcctcggatcagggcctgta 23084


Query: 4053  ctcctgcagggcggccaatcttcatggagagctgactgagagcacccagctgctgatcct 4112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 23085 ctcctgcagggcggccaatcttcatggagagctgactgagagcacccagctgctgatcct 23144


Query: 4113  ag 4114
              ||
Sbjct: 23145 ag 23146



 Score =  274 bits (138), Expect = 5e-70
 Identities = 138/138 (100%)
 Strand = Plus / Plus


Query: 4113  agatccccccccaagtccccacacagttggaagacatcagggccttgctcgctgccactgg 4172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 26524 agatccccccccaagtccccacacagttggaagacatcagggccttgctcgctgccactgg 26583


Query: 4173  accgaaccttccttcagtgctgacgtctcctctgggaacacagctggtcctggatcctgg 4232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 26584 accgaaccttccttcagtgctgacgtctcctctgggaacacagctggtcctggatcctgg 26643


Query: 4233  gaattctgctctccttgg 4250
              ||||||||||||||||||
Sbjct: 26644 gaattctgctctccttgg 26661



 Score =  264 bits (133), Expect = 5e-67
 Identities = 133/133 (100%)
 Strand = Plus / Plus


Query: 3803  caggaaagccactagtgaaaacgtcacgaatgacagtgatcaacacggagaagcctgcag 3862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 13789 caggaaagccactagtgaaaacgtcacgaatgacagtgatcaacacggagaagcctgcag 13848
```

```
Query: 3863  tcacagtcgatataggaagcaccatcaaaacagtgcagggagtgaatgtgacaatcaact 3922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 13849 tcacagtcgatataggaagcaccatcaaaacagtgcagggagtgaatgtgacaatcaact 13908


Query: 3923  gccaggttgcagg 3935
              |||||||||||||
Sbjct: 13909 gccaggttgcagg 13921
```

Score = 228 bits (115), Expect = 3e-56
Identities = 115/115 (100%)
Strand = Plus / Plus

```
Query: 4848   ccagtgcaatgggccttgcatcgggcctcacctagctgtgcaacacagacaagtcttctg 4907
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 102461 ccagtgcaatgggccttgcatcgggcctcacctagctgtgcaacacagacaagtcttctg 102520


Query: 4908   ccagacacgggatggcatcaccttaccatcagagcagtgcagtgctcttccgagg 4962
               |||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 102521 ccagacacgggatggcatcaccttaccatcagagcagtgcagtgctcttccgagg 102575
```

Score = 212 bits (107), Expect = 2e-51
Identities = 107/107 (100%)
Strand = Plus / Plus

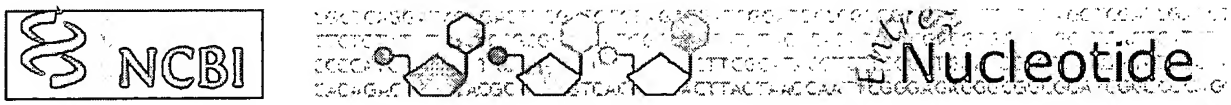
```
Query: 5183   tggagtgcagagacaccaccaggtactgcgagaaggtgaaacagctgaaactctgccaac 5242
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 105125 tggagtgcagagacaccaccaggtactgcgagaaggtgaaacagctgaaactctgccaac 105184


Query: 5243   tcagccagtttaaatctcgctgctgtggaacttgtggcaaagcgtga 5289
               |||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 105185 tcagccagtttaaatctcgctgctgtggaacttgtggcaaagcgtga 105231
```

☐ 1: AL591423. Human DNA sequenc...[gi:16973934] Links

```
LOCUS       AL591423               54193 bp    DNA     linear   PRI 16-NOV-2001
DEFINITION  Human DNA sequence from clone RP11-134P18 on chromosome 9, complete
            sequence.
ACCESSION   AL591423
VERSION     AL591423.6  GI:16973934
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 54193)
  AUTHORS   Almeida,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Nov 17, 2001 this sequence version replaced gi:16214807.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr9
            RP11-134P18 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            IMPORTANT: This sequence is not the entire insert of clone
            RP11-134P18 It may be shorter because we sequence overlapping
            sections only once, except for a short overlap.
            The true left end of clone RP11-220B22 is at 52194 in this
```

☐ 1: AL353895. Human DNA sequenc...[gi:13751339] Links

```
LOCUS       AL353895               163163 bp    DNA     linear   PRI 18-SEP-2001
DEFINITION  Human DNA sequence from clone RP11-503K16 on chromosome 9, complete
            sequence.
ACCESSION   AL353895
VERSION     AL353895.4  GI:13751339
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 163163)
  AUTHORS   Kimberley,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Apr 21, 2001 this sequence version replaced gi:13396472.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr9
            RP11-503K16 is from the library RPCI-11.2 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            This sequence is the entire insert of clone RP11-503K16 The true
            left end of clone RP11-134P18 is at 92104 in this sequence. The
            true right end of clone RP11-399M15 is at 92480 in this sequence.
FEATURES             Location/Qualifiers
```



1: AL158150. Human DNA sequenc...[gi:14160905]

Links

```
LOCUS       AL158150               168011 bp    DNA     linear   PRI 18-MAY-2001
DEFINITION  Human DNA sequence from clone RP11-220B22 on chromosome 9, complete
            sequence.
ACCESSION   AL158150
VERSION     AL158150.14  GI:14160905
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 168011)
  AUTHORS   Skuce,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-MAY-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On May 20, 2001 this sequence version replaced gi:13446402.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr9
            RP11-220B22 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            This sequence is the entire insert of clone RP11-220B22 The true
            left end of clone RP11-296P7 is at 58728 in this sequence.
FEATURES             Location/Qualifiers
     source          1..168011
```

☐ 1: AL442638. Homo sapiens chro...[gi:18857863] Links

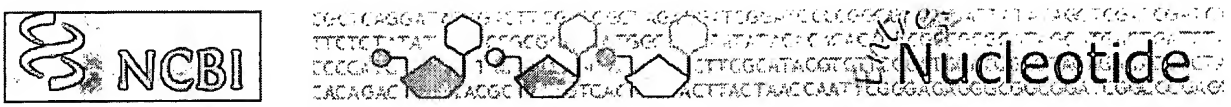
```
LOCUS       HS570H19              188247 bp    DNA     linear   PRI 19-FEB-2002
DEFINITION  Homo sapiens chromosome 9 BAC RP11-570H19, complete sequence.
ACCESSION   AL442638 AL358947
VERSION     AL442638.3  GI:18857863
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Plumb,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-AUG-2000) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
REFERENCE   2  (bases 1 to 188247)
  AUTHORS   Scharfe,M., Conrad,A., Hornischer,K., Loehnert,T.H., Thies,S. and
            Bloecker,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-SEP-2000) GBF, Dept. of Genome Analysis, Mascheroder
            Weg 1, D-38124 Braunschweig, Germany, E-mail: info.genome@gbf.de
COMMENT     On Feb 21, 2002 this sequence version replaced gi:11693452.
            All annotations in this database entry are developed by
            computational tools. It is therefore not explicitly noted in the
            feature lines that evidence is not experimental.
            Mapping was performed at The Sanger Centre
            (cf. http://www.sanger.ac.uk/HGP/Chr9)
            Mapping information is available via
            http://webace.sanger.ac.uk/cgi-bin/display?db=acedb9&grep=570H19
            -------------- Genome Center
            Center: GBF, Braunschweig
            Center code: GBF
            Web site: http://genome.gbf.de/
            Contact: info.genome@gbf.de
            -------------- Project Information
            Center project name:
            Center clone name: bA570H19
            -------------- Summary Statistics
            Sequencing vector: ###;
            Chemistry: Dye-terminator-BigDye: 65% of reads
            Chemistry: Dye-terminator-amersham: 31% of reads
            Chemistry: Dye-primer-amersham: 4% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 0 bases at least Q40
            Consensus quality: 0 bases at least Q30
            Consensus quality: 0 bases at least Q20
            Estimated insert size: ##; agarose-fp estimation
```

```
>NM_139238 ACCESSION:NM_139238 NID: gi 21327692 ref NM_139238.1
          Homo sapiens ADAMTS-like 1 (ADAMTSL1), transcript
          variant 1, mRNA
         Length = 2317

Identities = 668/669 (99%), Positives = 669/669 (100%)
 Frame = +2

Query: 1    MECCRRATPGTLLLFLAFLLLSSRTARSEEDRDGLWDAWGPWSECSRTCGGGASYSLRRC 60
            MECCRRATPGTLLLFLAFLLLSSRTARSEEDRDGLWDAWGPWSECSRTCGGGASYSLRRC
Sbjct: 65   MECCRRATPGTLLLFLAFLLLSSRTARSEEDRDGLWDAWGPWSECSRTCGGGASYSLRRC 244

Query: 61   LSSKSCEGRNIRYRTCSNVDCPPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNDPDNPCS 120
            LSSKSCEGRNIRYRTCSNVDCPPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNDPDNPCS
Sbjct: 245  LSSKSCEGRNIRYRTCSNVDCPPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNDPDNPCS 424

Query: 121  LKCQAKGTTLVVELAPKVLDGTRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNG 180
            LKCQAKGTTLVVELAPKVLDGTRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNG
Sbjct: 425  LKCQAKGTTLVVELAPKVLDGTRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNG 604

Query: 181  DGSTCRLVRGQYKSQLSATKSDDTVVAIPYGSRHIRLVLKGPDHLYLETKTLQGTKGENS 240
            DGSTCRLVRGQYKSQLSATKSDDTVVAIPYGSRHIRLVLKGPDHLYLETKTLQGTKGENS
Sbjct: 605  DGSTCRLVRGQYKSQLSATKSDDTVVAIPYGSRHIRLVLKGPDHLYLETKTLQGTKGENS 784

Query: 241  LSSTGTFLVDNSSVDFQKFPDKEILRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHR 300
            L+STGTFLVDNSSVDFQKFPDKEILRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHR
Sbjct: 785  LNSTGTFLVDNSSVDFQKFPDKEILRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHR 964

Query: 301  WRETDFFPCSATCGGGYQLTSAECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPCP 360
            WRETDFFPCSATCGGGYQLTSAECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPCP
Sbjct: 965  WRETDFFPCSATCGGGYQLTSAECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPCP 1144

Query: 361  ASDGYKQIMPYDLYHPLPRWEATPWTACSSSCGGGIQSRAVSCVEEDIQGHVTSVEEWKC 420
            ASDGYKQIMPYDLYHPLPRWEATPWTACSSSCGGGIQSRAVSCVEEDIQGHVTSVEEWKC
Sbjct: 1145ASDGYKQIMPYDLYHPLPRWEATPWTACSSSCGGGIQSRAVSCVEEDIQGHVTSVEEWKC 1324

Query: 421  MYTPKMPIAQPCNIFDCPKWLAQEWSPCTVTCGQGLRYRVVLCIDHRGMHTGGCSPKTKP 480
            MYTPKMPIAQPCNIFDCPKWLAQEWSPCTVTCGQGLRYRVVLCIDHRGMHTGGCSPKTKP
Sbjct: 1325MYTPKMPIAQPCNIFDCPKWLAQEWSPCTVTCGQGLRYRVVLCIDHRGMHTGGCSPKTKP 1504

Query: 481  HIKEECIVPTPCYKPKEKLPVEAKLPWFKQAQELEEGAAVSEEPSFIPEAWSACTVTCGV 540
            HIKEECIVPTPCYKPKEKLPVEAKLPWFKQAQELEEGAAVSEEPSFIPEAWSACTVTCGV
Sbjct: 1505HIKEECIVPTPCYKPKEKLPVEAKLPWFKQAQELEEGAAVSEEPSFIPEAWSACTVTCGV 1684

Query: 541  GTQVRIVRCQVLLSFSQSVADLPIDECEGPKPASQRACYAGPCSGEIPEFNPDETDGLFG 600
            GTQVRIVRCQVLLSFSQSVADLPIDECEGPKPASQRACYAGPCSGEIPEFNPDETDGLFG
Sbjct: 1685GTQVRIVRCQVLLSFSQSVADLPIDECEGPKPASQRACYAGPCSGEIPEFNPDETDGLFG 1864

Query: 601  GLQDFDELYDWEYEGFTKCSESCGGGVQEAVVSCLNKQTREPAEENLCVTSRRPPQLLKS 660
            GLQDFDELYDWEYEGFTKCSESCGGGVQEAVVSCLNKQTREPAEENLCVTSRRPPQLLKS
Sbjct: 1865GLQDFDELYDWEYEGFTKCSESCGGGVQEAVVSCLNKQTREPAEENLCVTSRRPPQLLKS 2044

Query: 661  CNLDPCPAR 669
            CNLDPCPAR
Sbjct: 2045CNLDPCPAR 2069
```

>NM_052866 ACCESSION:NM_052866 NID: gi 21327690 ref NM_052866.2
Homo sapiens ADAMTS-like 1 (ADAMTSL1), transcript
variant 2, mRNA
Length = 1810

Identities = 524/525 (99%), Positives = 525/525 (100%)
Frame = +2

```
Query: 1    MECCRRATPGTLLLFLAFLLLSSRTARSEEDRDGLWDAWGPWSECSRTCGGGASYSLRRC 60
            MECCRRATPGTLLLFLAFLLLSSRTARSEEDRDGLWDAWGPWSECSRTCGGGASYSLRRC
Sbjct: 65   MECCRRATPGTLLLFLAFLLLSSRTARSEEDRDGLWDAWGPWSECSRTCGGGASYSLRRC 244

Query: 61   LSSKSCEGRNIRYRTCSNVDCPPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNDPDNPCS 120
            LSSKSCEGRNIRYRTCSNVDCPPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNDPDNPCS
Sbjct: 245  LSSKSCEGRNIRYRTCSNVDCPPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNDPDNPCS 424

Query: 121  LKCQAKGTTLVVELAPKVLDGTRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNG 180
            LKCQAKGTTLVVELAPKVLDGTRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNG
Sbjct: 425  LKCQAKGTTLVVELAPKVLDGTRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNG 604

Query: 181  DGSTCRLVRGQYKSQLSATKSDDTVVAIPYGSRHIRLVLKGPDHLYLETKTLQGTKGENS 240
            DGSTCRLVRGQYKSQLSATKSDDTVVAIPYGSRHIRLVLKGPDHLYLETKTLQGTKGENS
Sbjct: 605  DGSTCRLVRGQYKSQLSATKSDDTVVAIPYGSRHIRLVLKGPDHLYLETKTLQGTKGENS 784

Query: 241  LSSTGTFLVDNSSVDFQKFPDKEILRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHR 300
            L+STGTFLVDNSSVDFQKFPDKEILRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHR
Sbjct: 785  LNSTGTFLVDNSSVDFQKFPDKEILRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHR 964

Query: 301  WRETDFFPCSATCGGGYQLTSAECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPCP 360
            WRETDFFPCSATCGGGYQLTSAECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPCP
Sbjct: 965  WRETDFFPCSATCGGGYQLTSAECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPCP 1144

Query: 361  ASDGYKQIMPYDLYHPLPRWEATPWTACSSSCGGGIQSRAVSCVEEDIQGHVTSVEEWKC 420
            ASDGYKQIMPYDLYHPLPRWEATPWTACSSSCGGGIQSRAVSCVEEDIQGHVTSVEEWKC
Sbjct: 1145ASDGYKQIMPYDLYHPLPRWEATPWTACSSSCGGGIQSRAVSCVEEDIQGHVTSVEEWKC 1324

Query: 421  MYTPKMPIAQPCNIFDCPKWLAQEWSPCTVTCGQGLRYRVVLCIDHRGMHTGGCSPKTKP 480
            MYTPKMPIAQPCNIFDCPKWLAQEWSPCTVTCGQGLRYRVVLCIDHRGMHTGGCSPKTKP
Sbjct: 1325MYTPKMPIAQPCNIFDCPKWLAQEWSPCTVTCGQGLRYRVVLCIDHRGMHTGGCSPKTKP 1504

Query: 481  HIKEECIVPTPCYKPKEKLPVEAKLPWFKQAQELEEGAAVSEEPS 525
            HIKEECIVPTPCYKPKEKLPVEAKLPWFKQAQELEEGAAVSEEPS
Sbjct: 1505HIKEECIVPTPCYKPKEKLPVEAKLPWFKQAQELEEGAAVSEEPS 1639
```



☐ 1: NM_139238. Homo sapiens ADAM...[gi:21327692] Links

```
LOCUS       NM_139238               2317 bp    mRNA    linear   PRI 07-MAY-2003
DEFINITION  Homo sapiens ADAMTS-like 1 (ADAMTSL1), transcript variant 1, mRNA.
ACCESSION   NM_139238
VERSION     NM_139238.1  GI:21327692
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2317)
  AUTHORS   Hirohata,S., Wang,L.W., Miyagi,M., Yan,L., Seldin,M.F., Keene,D.R.,
            Crabb,J.W. and Apte,S.S.
  TITLE     Punctin, a novel ADAMTS-like molecule, ADAMTSL-1, in extracellular
            matrix
  JOURNAL   J. Biol. Chem. 277 (14), 12182-12189 (2002)
  MEDLINE   21922817
   PUBMED   11805097
  REMARK    GeneRIF: Punctin, a novel ADAMTS-like molecule, ADAMTSL-1, in
            extracellular matrix
COMMENT     REVIEWED REFSEQ: This record has been curated by NCBI staff. The
            reference sequence was derived from AF251058.1 and BC030262.1.

            Summary: This gene encodes a secreted protein resembling members of
            the ADAMTS (a disintegrin and metalloproteinase with thrombospondin
            motif) family. This protein lacks the propeptide region and the
            metalloproteinase and disintegrin-like domains, which are typical
            of the ADAMTS family, but contains other ADAMTS domains, including
            the thrombospondin type 1 motif. This protein may have important
            functions in the extracellular matrix. Alternative splicing of this
            gene results in 3 transcript variants encoding different isoforms.

            Transcript Variant: This variant (1) encodes the longest isoform
            (1).
FEATURES             Location/Qualifiers
     source          1..2317
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="9"
                     /map="9p22.1"
     gene            1..2317
                     /gene="ADAMTSL1"
                     /note="synonyms: ADAMTSR1, MGC40193"
                     /db_xref="LocusID:92949"
     CDS             65..2116
                     /gene="ADAMTSL1"
                     /note="ADAM-TS related protein 1; thrombospondin; punctin"
```

```
                     /codon_start=1
                     /product="ADAM-TS related protein 1 isoform 1"
                     /protein_id="NP_640329.1"
                     /db_xref="GI:21327693"
                     /db_xref="LocusID:92949"
                     /translation="MECCRRATPGTLLLFLAFLLLSSRTARSEEDRDGLWDAWGPWSE
                     CSRTCGGGASYSLRRCLSSKSCEGRNIRYRTCSNVDCPPEAGDFRAQQCSAHNDVKHH
                     GQFYEWLPVSNDPDNPCSLKCQAKGTTLVVELAPKVLDGTRCYTESLDMCISGLCQIV
                     GCDHQLGSTVKEDNCGVCNGDGSTCRLVRGQYKSQLSATKSDDTVVAIPYGSRHIRLV
                     LKGPDHLYLETKTLQGTKGENSLNSTGTFLVDNSSVDFQKFPDKEILRMAGPLTADFI
                     VKIRNSGSADSTVQFIFYQPIIHRWRETDFFPCSATCGGGYQLTSAECYDLRSNRVVA
                     DQYCHYYPENIKPKPKLQECNLDPCPASDGYKQIMPYDLYHPLPRWEATPWTACSSSC
                     GGGIQSRAVSCVEEDIQGHVTSVEEWKCMYTPKMPIAQPCNIFDCPKWLAQEWSPCTV
                     TCGQGLRYRVVLCIDHRGMHTGGCSPKTKPHIKEECIVPTPCYKPKEKLPVEAKLPWF
                     KQAQELEEGAAVSEEPSFIPEAWSACTVTCGVGTQVRIVRCQVLLSFSQSVADLPIDE
                     CEGPKPASQRACYAGPCSGEIPEFNPDETDGLFGGLQDFDELYDWEYEGFTKCSESCG
                     GGVQEAVVSCLNKQTREPAEENLCVTSRRPPQLLKSCNLDPCPARSSIDSAWNACNVL
                     C"
     misc_feature    170..310
                     /gene="ADAMTSL1"
                     /note="TSP1; Region: Thrombospondin type 1 repeats"
                     /db_xref="CDD:smart00209"
     variation       complement(217)
                     /allele="T"
                     /allele="A"
                     /db_xref="dbSNP:2277160"
BASE COUNT      554 a    619 c    619 g    525 t
ORIGIN
        1 gcaggcagag gagcacttag cagcttattc agtgtccgat tctgattccg gcaaggatcc
       61 aagcatggaa tgctgccgtc gggcaactcc tggcacactg ctcctctttc tggctttcct
      121 gctcctgagt tccaggaccg cacgctccga ggaggaccgg gacggcctat gggatgcctg
      181 gggcccatgg agtgaatgct cacgcacctg cgggggtggg gcctcctact ctctgaggcg
      241 ctgcctgagc agcaagagct gtgaaggaag aaatatccga tacagaacat gcagtaatgt
      301 ggactgccca ccagaagcag gtgatttccg agctcagcaa tgctcagctc ataatgatgt
      361 caagcaccat ggccagtttt atgaatggct tcctgtgtct aatgaccctg acaacccatg
      421 ttcactcaag tgccaagcca aaggaacaac cctggttgtt gaactagcac ctaaggtctt
      481 agatggtacg cgttgctata cagaatcttt ggatatgtgc atcagtggtt tatgccaaat
      541 tgttggctgc gatcaccagc tgggaagcac cgtcaaggaa gataactgtg gggtctgcaa
      601 cggagatggg tccacctgcc ggctggtccg agggcagtat aaatcccagc tctccgcaac
      661 caaatcggat gatactgtgg ttgcaattcc ctatggaagt agacatattc gccttgtctt
      721 aaaaggtcct gatcacttat atctggaaac caaaaccctc caggggacta aaggtgaaaa
      781 cagtctcaac tccacaggaa ctttccttgt ggacaattct agtgtggact tccagaaatt
      841 tccagacaaa gagatactga gaatggctgg accactcaca gcagatttca ttgtcaagat
      901 tcgtaactcg ggctccgctg acagtacagt ccagttcatc ttctatcaac ccatcatcca
      961 ccgatggagg gagacggatt tctttccttg ctcagcaacc tgtggaggag gttatcagct
     1021 gacatcggct gagtgctacg atctgaggag caaccgtgtg gttgctgacc aatactgtca
     1081 ctattaccca gagaacatca aacccaaacc caagcttcag gagtgcaact tggatccttg
     1141 tccagccagt gacggataca agcagatcat gccttatgac ctctaccatc cccttcctcg
     1201 gtgggaggcc accccatgga ccgcgtgctc ctcctcgtgt gggggggca tccagagccg
     1261 ggcagtttcc tgtgtggagg aggacatcca ggggcatgtc acttcagtgg aagagtggaa
     1321 atgcatgtac acccctaaga tgcccatcgc gcagccctgc aacatttttg actgccctaa
     1381 atggctggca caggagtggt ctccgtgcac agtgacatgt ggccagggcc tcagataccg
     1441 tgtggtcctc tgcatcgacc atcgaggaat gcacacagga ggctgtagcc caaaaacaaa
     1501 gccccacata aaagaggaat gcatcgtacc cactccctgc tataaaccca aagagaaact
     1561 tccagtcgag gccaagttgc catggttcaa acaagctcaa gagctagaag aaggagctgc
     1621 tgtgtcagag gagccctcgt tcatcccaga ggcctggtcg gcctgcacag tcacctgtgg
     1681 tgtggggacc caggtgcgaa tagtcaggtg ccaggtgctc ctgtctttct ctcagtccgt
     1741 ggctgacctg cctattgacg agtgtgaagg gcccaagcca gcatcccagc gtgcctgtta
     1801 tgcaggccca tgcagcgggg aaattcctga gttcaaccca gacgagacag atgggctctt
     1861 tggtggcctg caggatttcg acgagctgta tgactgggag tatgaggggt tcaccaagtg
```

```
     1921 ctccgagtcc tgtggaggag gtgtccagga ggctgtggtg agctgcttga acaaacagac
     1981 tcgggagcct gctgaggaga acctgtgcgt gaccagccgc cggcccccac agctcctgaa
     2041 gtcctgcaat ttggatccct gcccagcaag aagcagtatc gactcagcat ggaacgcctg
     2101 caacgttctt tgttaggcaa ccaagaggcc tggcttctca tcctgctgtc accaactagc
     2161 tctgtggcct agggcgaggt gtctgccctt tatgtttcca catctgcaaa gtgaactggt
     2221 tgtacctgat gatctgagat cccatgactt gctcacatgt cccatgattc tttatttttgt
     2281 aggcagaagc attaaacagc tactcctgct gctgtgt
//
```


Jun 19 2003 12:37:45

☐ **1:** NM_052866. Homo sapiens ADAM...[gi:21327690] Links

```
LOCUS       NM_052866               1810 bp    mRNA    linear   PRI 07-MAY-2003
DEFINITION  Homo sapiens ADAMTS-like 1 (ADAMTSL1), transcript variant 2, mRNA.
ACCESSION   NM_052866
VERSION     NM_052866.2  GI:21327690
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1810)
  AUTHORS   Hirohata,S., Wang,L.W., Miyagi,M., Yan,L., Seldin,M.F., Keene,D.R.,
            Crabb,J.W. and Apte,S.S.
  TITLE     Punctin, a novel ADAMTS-like molecule, ADAMTSL-1, in extracellular
            matrix
  JOURNAL   J. Biol. Chem. 277 (14), 12182-12189 (2002)
  MEDLINE   21922817
   PUBMED   11805097
  REMARK    GeneRIF: Punctin, a novel ADAMTS-like molecule, ADAMTSL-1, in
            extracellular matrix
COMMENT     REVIEWED REFSEQ: This record has been curated by NCBI staff. The
            reference sequence was derived from AF176313.1 and BC030262.1.
            On Jun 6, 2002 this sequence version replaced gi:16418368.

            Summary: This gene encodes a secreted protein resembling members of
            the ADAMTS (a disintegrin and metalloproteinase with thrombospondin
            motif) family. This protein lacks the propeptide region and the
            metalloproteinase and disintegrin-like domains, which are typical
            of the ADAMTS family, but contains other ADAMTS domains, including
            the thrombospondin type 1 motif. This protein may have important
            functions in the extracellular matrix. Alternative splicing of this
            gene results in 3 transcript variants encoding different isoforms.

            Transcript Variant: This variant (2) has alternate 3' exons, as
            compared to variant 1, resulting in immediate translation
            termination. Isoform 2 is truncated at the C-terminus, compared to
            isoform 1.
            COMPLETENESS: complete on the 3' end.
FEATURES             Location/Qualifiers
     source          1..1810
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="9"
                     /map="9p22.1"
     gene            1..1810
                     /gene="ADAMTSL1"
                     /note="synonyms: ADAMTSR1, MGC40193"
```

```
                     /db_xref="LocusID:92949"
     CDS             65..1642
                     /gene="ADAMTSL1"
                     /note="ADAM-TS related protein 1; thrombospondin; punctin"
                     /codon_start=1
                     /product="ADAM-TS related protein 1 isoform 2"
                     /protein_id="NP_443098.2"
                     /db_xref="GI:21327691"
                     /db_xref="LocusID:92949"
                     /translation="MECCRRATPGTLLLFLAFLLLSSRTARSEEDRDGLWDAWGPWSE
                     CSRTCGGGASYSLRRCLSSKSCEGRNIRYRTCSNVDCPPEAGDFRAQQCSAHNDVKHH
                     GQFYEWLPVSNDPDNPCSLKCQAKGTTLVVELAPKVLDGTRCYTESLDMCISGLCQIV
                     GCDHQLGSTVKEDNCGVCNGDGSTCRLVRGQYKSQLSATKSDDTVVAIPYGSRHIRLV
                     LKGPDHLYLETKTLQGTKGENSLNSTGTFLVDNSSVDFQKFPDKEILRMAGPLTADFI
                     VKIRNSGSADSTVQFIFYQPIIHRWRETDFFPCSATCGGGYQLTSAECYDLRSNRVVA
                     DQYCHYYPENIKPKPKLQECNLDPCPASDGYKQIMPYDLYHPLPRWEATPWTACSSSC
                     GGGIQSRAVSCVEEDIQGHVTSVEEWKCMYTPKMPIAQPCNIFDCPKWLAQEWSPCTV
                     TCGQGLRYRVVLCIDHRGMHTGGCSPKTKPHIKEECIVPTPCYKPKEKLPVEAKLPWF
                     KQAQELEEGAAVSEEPS"
     misc_feature    170..310
                     /gene="ADAMTSL1"
                     /note="TSP1; Region: Thrombospondin type 1 repeats"
                     /db_xref="CDD:smart00209"
     variation       complement(217)
                     /allele="T"
                     /allele="A"
                     /db_xref="dbSNP:2277160"
     variation       1678..1679
                     /gene="ADAMTSL1"
                     /allele="GT"
                     /allele="-"
                     /db_xref="dbSNP:3833713"
     polyA_signal    1771..1776
                     /gene="ADAMTSL1"
     polyA_site      1795
                     /gene="ADAMTSL1"
BASE COUNT      481 a    459 c    453 g    417 t
ORIGIN
        1 gcaggcagag gagcacttag cagcttattc agtgtccgat tctgattccg gcaaggatcc
       61 aagcatggaa tgctgccgtc gggcaactcc tggcacactg ctcctctttc tggctttcct
      121 gctcctgagt tccaggaccg cacgctccga ggaggaccgg gacggcctat gggatgcctg
      181 gggcccatgg agtgaatgct cacgcacctg cgggggtggg gcctcctact ctctgaggcg
      241 ctgcctgagc agcaagagct gtgaaggaag aaatatccga tacagaacat gcagtaatgt
      301 ggactgccca ccagaagcag gtgatttccg agctcagcaa tgctcagctc ataatgatgt
      361 caagcaccat ggccagtttt atgaatggct tcctgtgtct aatgaccctg acaacccatg
      421 ttcactcaag tgccaagcca aaggaacaac cctggttgtt gaactagcac ctaaggtctt
      481 agatggtacg cgttgctata cagaatcttt ggatatgtgc atcagtggtt tatgccaaat
      541 tgttggctgc gatcaccagc tgggaagcac cgtcaaggaa gataactgtg gggtctgcaa
      601 cggagatggg tccacctgcc ggctggtccg agggcagtat aaatcccagc tctccgcaac
      661 caaatcggat gatactgtgg ttgcaattcc ctatggaagt agacatattc gccttgtctt
      721 aaaaggtcct gatcacttat atctggaaac caaaaccctc caggggacta aaggtgaaaa
      781 cagtctcaac tccacaggaa ctttccttgt ggacaattct agtgtggact tccagaaatt
      841 tccagacaaa gagatactga gaatggctgg accactcaca gcagatttca ttgtcaagat
      901 tcgtaactcg ggctccgctg acagtacagt ccagttcatc ttctatcaac ccatcatcca
      961 ccgatggagg gagacggatt tctttccttg ctcagcaacc tgtggaggag gttatcagct
     1021 gacatcggct gagtgctacg atctgaggag caaccgtgtg gttgctgacc aatactgtca
     1081 ctattaccca gagaacatca aacccaaacc caagcttcag gagtgcaact tggatccttg
     1141 tccagccagt gacggataca agcagatcat gccttatgac ctctaccatc cccttcctcg
     1201 gtgggaggcc accccatgga ccgcgtgctc ctcctcgtgt gggggggggca tccagagccg
     1261 ggcagtttcc tgtgtggagg aggacatcca ggggcatgtc acttcagtgg aagagtggaa
```

```
     1321 atgcatgtac acccctaaga tgcccatcgc gcagccctgc aacatttttg actgccctaa
     1381 atggctggca caggagtggt ctccgtgcac agtgacatgt ggccagggcc tcagataccg
     1441 tgtggtcctc tgcatcgacc atcgaggaat gcacacagga ggctgtagcc caaaaacaaa
     1501 gccccacata aaagaggaat gcatcgtacc cactccctgc tataaaccca aagagaaact
     1561 tccagtcgag gccaagttgc catggttcaa acaagctcaa gagctagaag aaggagctgc
     1621 tgtgtcagag gagccctcgt aagttgtaaa agcacagact gttctatatt tgaaactgtt
     1681 ttgtttaaag aaagcagtgt ctcactggtt gtagctttca tgggttctga actaagtgta
     1741 atcatctcac caaagctttt tggctctcaa attaaagatt gattagtttc aaaaaaaaaa
     1801 aaaaaaaaaa
//
```

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 277, No. 14, Issue of April 5, pp. 12182–12189, 2002
Printed in U.S.A.

Punctin, a Novel ADAMTS-like Molecule, ADAMTSL-1, in Extracellular Matrix*

Received for publication, October 5, 2001, and in revised form, January 17, 2002
Published, JBC Papers in Press, January 22, 2002, DOI 10.1074/jbc.M109665200

Satoshi Hirohata‡§, Lauren W. Wang‡§, Masaru Miyagi¶, Lin Yan¶, Michael F. Seldin‖, Douglas R. Keene, John W. Crabb¶, and Suneel S. Apte‡ ‡§§**

*From the ‡Department of Biomedical Engineering, Lerner Research Institute, the ‡‡Department of Orthopaedic Surgery, Cleveland Clinic Foundation, the ¶Department of Ophthalmic Research, Cole Eye Institute, Cleveland Clinic Foundation, Cleveland, Ohio 44195, the ‖Rowe Program in Genetics, Departments of Biological Chemistry and Medicine, University of California at Davis, Davis, California 956161, and the **Shriner's Hospital for Children, Portland, Oregon 97201*

Punctin (ADAMTSL-1) is a secreted molecule resembling members of the ADAMTS family of proteases. Punctin lacks the pro-metalloprotease and the disintegrin-like domain typical of this family but contains other ADAMTS domains in precise order including four thrombospondin type I repeats. Punctin is the product of a distinct gene on human chromosome 9p21-22 and mouse chromosome 4 that is expressed in adult skeletal muscle. His-tagged punctin expressed in stably transfected High-Five™ insect cells was purified to apparent homogeneity by Ni-chromatography of conditioned medium. The NH_2 terminus is not blocked and has the sequence EEDRD and so forth as determined by Edman degradation, demonstrating signal peptidase processing. Recombinant epitope-tagged punctin has a calculated mass of 59,991 Da but exhibits major molecular species of 61970 ± 6 Da and 62131 ± 5 Da as measured by liquid chromatography electrospray mass spectrometry. Punctin is a glycoprotein based on carbohydrate staining and liquid chromatography electrospray mass spectrometry glycopeptide analysis. Glycosylation occurs at a single *N*-linked site as demonstrated by altered electrophoretic migration of punctin expressed in the presence of tunicamycin A. Punctin contains disulfide bonds based on antibody accessibility and electrophoretic migration under reducing *versus* nonreducing conditions. Rotary shadowing demonstrates that punctin is hatchet-shaped having a globular region attached to a short stem. In transfected COS-1 cells, punctin is deposited in the cell substratum in a punctate fashion and is excluded from focal contacts. Punctin is the first member of a novel family of ADAMTS-like proteins that may have important functions in the extracellular matrix.

Metalloproteases responsible for extracellular (ECM)[1] turnover have a modular structure. Matrix metalloproteinases (MMPs) (1), a disintegrin-like and metalloprotease (ADAMs) (2), and proteases of the ADAMTS family (3, 4) are composed of characteristic domains arranged in a precise order that is the hallmark of each family. These enzymes are structurally and functionally bipartite consisting of an enzymatic domain attached to nonenzymatic or ancillary domains. The ancillary domains localize these proteases to substrates, the cell surface, or to the ECM. The ancillary domains of the gelatinases MMP-2 and MMP-9 are among the best studied of the substrate-binding domains. The fibronectin type II domains of the gelatinases are involved in binding to gelatin and some collagens as well as to fibronectin and heparin as in the case of MMP-2 (5, 6). The gelatin-binding domain of MMP-2 binds the matricellular proteins thrombospondin-1 (TSP1) and TSP2 (7). Although neither is a substrate for MMP-2, the interaction may mediate the clearance of MMP-2 and affect cell-adhesive properties (8). The MMP-2 hemopexin domain interacts with the carboxyl terminus of the tissue inhibitor of metalloproteases-2, facilitating pro-MMP-2 activation by membrane-type MMPs (1, 5, 6, 9). The MMP-2 hemopexin domain also interacts with a chemokine called monocyte chemoattractant protein-3, which allows its processing by the catalytic domain (10). The disintegrin domains of ADAMs such as ADAM-15 are implicated in cell-cell adhesion (2, 11, 12), and the ancillary domains of ADAMTS-1 are required for its binding to the ECM (13). In some ADAMs, the zinc-binding active site is nonfunctional, suggesting that they do not function as proteases at all but may instead have a primary role in adhesion via their ancillary domains (2).

With this background, it is conceptually possible that gene products containing only the ancillary domains of ADAMTS may have specific functions in cell-cell or cell-matrix interactions or may regulate ADAMTS proteases. We have identified an ADAMTS-like (ADAMTSL) molecule named punctin,[2]

* This work was supported in part by the Cleveland Clinic Foundation (to S. S. A.), a Yamanouchi USA Foundation Award (to S. S. A.), and National Institutes of Health Grants EY06603 (to J. W. C.) and HGO0734 (to M. F. S.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AF176313.

§ Both authors contributed equally to this work.

§§ To whom correspondence should be addressed: Dept. of Biomedical Engineering (ND20), Lerner Research Inst., Cleveland Clinic Foundation, 9500 Euclid Ave., Cleveland, OH 44195. Tel.: 216-445-3278; Fax: 216-445-4383; E-mail: aptes@bme.ri.ccf.org or www.lerner.ccf.org/pi/apte.html.

[1] The abbreviations used are: ECM, extracellular matrix; ADAMTSL, a disintegrin-like and metalloprotease domain with thrombospondin type I motifs like; ADAMTS, a disintegrin-like and metalloprotease domain with thrombospondin type I motifs; ADAM, a disintegrin-like and metalloprotease; MS, mass spectrometry; EST, expressed sequence tag; LC-ESMS, liquid chromatography-electrospray mass spectrometry; MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight; MMP, matrix metalloprotease; ORF, open reading frame; PBS, phosphate-buffered saline; RACE, rapid amplification of cDNA ends; TSP, thrombospondin; TS, thrombospondin type I domain; HexNAc, *N*-acetylhexosamine; NeuAc, *N*-acetylneuraminic acid.

[2] Approved gene symbols *ADAMTSL1* and *Adamtsl1* indicate human and mouse orthologs, respectively. The corresponding protein product of these genes, ADAMTSL-1, is designated by the trivial name punctin because of its punctate distribution beneath transfected cells.

which is the product of a gene distinct from any in the ADAMTS family and is composed of ADAMTS ancillary domains alone. We have purified and characterized recombinant punctin produced in insect cells, visualized it by electron microscopy, and demonstrated that it is a glycoprotein and a component of the ECM.

EXPERIMENTAL PROCEDURES

cDNA Cloning and Sequence Analysis—Using BLAST programs from the National Center for Biotechnology Information, we scanned the data base of ESTs using the protein sequences of ADAMTS proteases previously cloned by us (4, 14) and identified a human EST (GenBank™ accession number AA482392 encoded by IMAGE clone 752797). The EST predicted a polypeptide with a similarity to the carboxyl half of cognate ADAMTS members but with no identities in GenBank™ or other protein and nucleotide data bases.

Using nested oligonucleotide primers based on the sequences at the 5′ and 3′ ends of the IMAGE clone insert and human skeletal muscle cDNA (Marathon cDNA, CLONTECH, Palo Alto, CA) as the template, we performed RACE and extended the cDNA at 5′ and 3′ ends by PCR essentially as described previously (4, 14).

Northern Blot Analysis—Multiple tissue Northern blots from adult human and mouse tissues (CLONTECH, Palo Alto, CA) were hybridized to a [α-^{32}P]dCTP-labeled punctin probe, a 1200-bp cDNA fragment from the 5′ end of the punctin coding sequence, followed by autoradiographic exposure for 7 days.

Chromosomal Mapping and Genomic Arrangement—To determine the chromosomal location of ***Adamtsl1***, we analyzed a panel of DNA samples from an interspecific cross that has been characterized for over 1200 genetic markers throughout the mouse genome (15). Markers can be seen on the worldwide web (www.informatics.jax.org/searches/cross-data_form.shtml) by entering "DNA Mapping Panel Data Sets" from the mouse genome data base and then selecting the "Seldin cross" and "Chromosome." Initially, DNA from the two parental mice, (C3H/HeJ-*gld*) and (C3H/HeJ-*gld* × *Mus spretus*) F_1), were digested with various restriction endonucleases and hybridized with the ***Adamtsl1*** cDNA probe (IMAGE clone 2076907 with GenBank™ accession number AI787975) to determine restriction fragment length variants for haplotype analyses. Gene linkage was determined by segregation analysis. Gene order was determined by analyzing all haplotypes and minimizing crossover frequency among all genes that were determined to be within a linkage group. This method resulted in the determination of the most probable gene order. To define the locus for *ADAMTSL1*, the human punctin cDNA sequence was used for BLAST searches of the human genome (Celera Sciences, Rockville, MD).

Generation and Characterization of Anti-punctin Antisera—The peptide (NH_2)-[C]YYPENIKPKPKLQE-(OH) located in the third TS domain of punctin (Fig. 1*B*) was synthesized using Fmoc (*N*-(9-fluorenyl) methoxycarbonyl) chemistry, purified by reverse-phase high-pressure liquid chromatography, and molecular weight was confirmed by MS (Alpha Diagnostic International, San Antonio, TX). A cysteine ([C]) residue was included at the NH_2 terminus for coupling to keyhole limpet hemocyanin. Peptide-keyhole limpet hemocyanin conjugate was dialyzed in PBS and used for immunization. Two New Zealand White male rabbits (7–8 pounds) were immunized with the conjugate (~200 μg/injection/rabbit, multiple intramuscular and subcutaneous sites) at biweekly intervals for 8 weeks. After an initial injection in Freund's complete adjuvant, subsequent injections were given in incomplete adjuvant. Antibody titer was measured by enzyme-linked immunosorbent assay using free peptide.

Immune sera were tested by Western blot analysis of extracts from COS-1 cells transiently transfected with punctin cDNA (see below). Although antisera from both rabbits (antisera 4112 and 4113) gave qualitatively similar results, the best signal/noise ratio was obtained with antiserum 4113. Affinity-purified antibodies were prepared by column chromatography of antiserum 4113 using the immobilized peptide immunogen.

Expression and Purification of Recombinant Punctin from Insect Cells—High-Five™ cells (Invitrogen) were routinely cultured on tissue culture plastic and maintained at 27 °C in Ultimate™ serum-free insect cell medium (Invitrogen) as per manufacturer's directions. The full-length punctin ORF was excised from pcDNA3.1/Myc-His B-TSL1 (see below) with *Eco*RI and *Not*I and ligated into the corresponding sites in pIZT/V5-His (Invitrogen). The resulting insect cell expression plasmid pIZT/V5-His-TSL1 generated punctin with a COOH-terminal V5 epitope and 6× His tag. pIZT/V5-His-TSL1 was transfected into High-Five™ cells using Insectin-Plus liposomes (Invitrogen) and plated onto 100-mm Petri dishes. After 48 h, antibiotic selection (500 μg/ml Zeocin, Invitrogen) was started and continued for 21 days. Colonies that survived selection were picked manually, expanded, and maintained in medium containing Zeocin (50 μg/ml). Punctin production by isolated colonies was tested by Western blot analysis of conditioned medium using anti-His monoclonal antibody (Invitrogen) and antibody 4113.

For protein production, cells were grown in suspension in either Ultimate™ serum-free insect cell medium or Express-Five serum-free medium containing heparin (5 units/ml, Invitrogen). Production cultures were in spinner flasks, and culture medium was stored at −80 °C with 1 mM phenylmethylsulfonyl fluoride until use. For purification, medium was dialyzed into binding buffer (20 mM sodium phosphate, 500 mM NaCl, pH 7.8) containing 0.03% Brij-35 (Sigma). Purification was performed using 1-liter batches of dialyzed medium and a 5-ml Ni-Sepharose column (ProBond™, Invitrogen) on an fast protein liquid chromatography instrument (Bio-Rad, Hercules, CA). Following binding, the column was washed with three column volumes of binding buffer. A gradient of 0–42.5 mM imidazole in binding buffer was used to remove nonspecifically bound molecules from the column. Elution was with four column volumes of 250 mM imidazole in binding buffer, pH 7.0, containing 0.03% Brij-35. Elution was monitored by in-line UV and conductivity measurements. 2-ml fractions of eluate were collected and tested by Western blot analysis as described above. Fractions containing punctin were pooled. Protein concentration was determined using the Bradford assay (Bio-Rad) and by phenylthiocarbamyl amino acid analysis using an Applied Biosystems model 420H/130/920 automated analysis system (16).

Characterization of Recombinant Punctin—The NH_2-terminal sequence of recombinant punctin was determined by Edman degradation. Recombinant punctin (5 μg) was electrophoresed on 10% SDS-PAGE, electrotransferred to polyvinylidene difluoride membrane, and lightly stained with modified Coomassie Blue (Simply Blue Safe Stain, Invitrogen). Protein bands were excised and subjected to Edman degradation on an Applied Biosystems Procise 492 sequencer in the Molecular Biotechnology Core Facility of the Lerner Research Institute.

To probe for glycosylation, recombinant punctin (4 μg) was electrophoresed on 10% SDS-PAGE and stained for carbohydrate using a periodic acid-Schiff reaction-based method (Pro-Q fuchsia glycoprotein staining kit, Molecular Probes, Eugene, OR). In this reaction, CandyCane™ glycoprotein molecular weight standards consisting of alternate bands of glycosylated and unglycosylated proteins were used as controls. Glycoprotein staining was also performed after enzymatic deglycosylation of punctin with peptide *N*-glycosidase F. Deglycosylation of denatured as well as native punctin was performed with a commercially available kit (Bio-Rad) using bovine fetuin as a control. To investigate further whether *N*-linked carbohydrates were present in punctin, stably transfected insect cells were cultured in the presence or absence of tunicamycin A1 homolog (0.1 μg/ml culture medium, Sigma). Equal amounts of total protein from culture medium of tunicamycin-treated and untreated cells were assayed by Western blot with antibody 4113 at various time points after the addition of tunicamycin

Mass Spectrometry—The molecular mass of punctin was measured by MALDI-TOF and by LC-ESMS. MALDI-TOF was performed with a PerkinElmer Biosystems Voyager DE Pro-mass spectrometer using sinapinic acid as the matrix and bovine serum albumin as a calibration standard protein (17). MALDI-TOF MS measurements of intact punctin and naturally observed limited proteolysis fragments are reported ± 50% peak width (in Da) at half-maximal peak height. LC-ESMS was performed with a PerkinElmer Sciex API 3000 triple quadruple mass spectrometer (17, 18). Nitrogen was used as the nebulization gas at 40 p.s.i., and curtain gas was supplied from a nitrogen generator (Whatman model 75-72). For LC-ESMS of intact punctin, a scan range of 700–1800 *m/z* was used with 0.2 atomic mass unit steps, a scan time of 7.5 s, and at an orifice potential of 80 and 5000 V ion spray. Reverse phase-high-pressure liquid chromatography was done at a flow rate of 5 μl/min on a 5-μm Vydac C18 capillary column (0.3 × 150 mm, LC Packing) using an Applied Biosystems Model 140D high-pressure liquid chromatography system and aqueous acetonitrile/trifluoroacetic acid solvents with 100% of the eluant going to the mass spectrometer. ESMS measurements of intact punctin are reported as the mean ± S.E. (in Da).

For glycopeptide characterization, punctin was excised from a SDS-polyacrylamide gel (~1 μg/lane × 6 lanes), in-gel reduced with 10 mM dithiothreitol, cysteine-alkylated with 20 mM iodoacetamide in 400 mM ammonium bicarbonate, and digested with 0.2 μg of trypsin (Promega) overnight at 37 °C in 100 mM ammonium bicarbonate. Peptides from the in-gel tryptic digests were extracted with 60% acetonitrile containing 0.1% trifluoroacetic acid, dried in a Speed Vac, redissolved in 50 μl

of 0.1% trifluoroacetic acid, and analyzed by LC-ESMS using selective ion monitoring with the PE Sciex API 3000 triple quadruple mass spectrometer system as described above for intact protein analyses. Glycopeptides were selectively detected based on diagnostic sugar oxonium ions HexNAc + Hex (*m/z* 366) and *N*-acetylneuraminic acid (NeuAc) (*m/z* 292) (17). Carbohydrate marker ions at *m/z* 366 and 292 (dwell time 200 ms each) were monitored in a positive ion mode at a high orifice potential (180 V), whereas full scans at *m/z* 300–2300 (0.2 atomic mass unit steps, scan time 3.5 s) were acquired at a lower orifice potential (70 V). This way both intact parent ions and abundant marker ions were observed in the same *m/z* scan.

Rotary Shadowing and Electron Microscopy of Recombinant Punctin—Rotary shadowing was done essentially as described previously (19). A 30-μl sample of punctin at 100 μg/ml was mixed with 70 μl of glycerol and nebulized onto freshly cleaved mica using an airbrush. The sample was dried in a vacuum, and rotary shadowed using a platinum-carbon electron beam gun angled at 6° relative to the mica surface within a Balzers BAE 250 evaporator. The replica was backed with carbon, floated onto distilled water, and picked up onto 600 mesh grids. Photomicrographs were taken using a Philips 410 electron microscope operated at 80 kV.

Transient Expression of Tagged and Untagged Punctin in COS-1 Cells—An internal *Sac*I site and a flanking *Not*I site were used to remove a 1.5-kb fragment of IMAGE clone 752797 and ligate it into corresponding sites in IMAGE clone 2150669 corresponding to the 5′ end of the punctin cDNA to generate a complete ORF. *Eco*RI and *Not*I sites flanking this ORF were used to excise and clone the full-length coding sequence into pcDNA3.1/Myc-His (+) A (Invitrogen) for the expression of untagged punctin. To make constructs in which the *ADAMTSL1* ORF was in-frame with a carboxyl-terminal FLAG tag or a tandem *myc* tag and 6× His tag, PCR was performed with Advantage 2 polymerase (CLONTECH, Palo Alto, CA) using the full-length coding sequence as a template. The amplicons were cloned into the vectors pFLAG-CMV5c (Sigma) and pcDNA3.1/Myc-His B (Invitrogen) for expression with either a COOH-terminal FLAG tag or a COOH-terminal tandem *myc* tag and 6× His tag, respectively.

COS-1 cells (ATCC number CRL-1650) were grown on tissue culture plastic in Dulbecco's modified Eagle's medium:F-12 (1:1) (Lerner Research Institute Media Services) supplemented with 10% fetal bovine serum (Invitrogen) and antibiotics (100 units/ml of penicillin and 50 μg/ml streptomycin). 10^5 cells between passages 3 and 10 were transfected with untagged, FLAG-tagged, or *myc* + 6× His-tagged punctin using FuGENE 6 (Roche Molecular Biochemicals) as per manufacturer's recommendations, and cells were grown for an additional 24–48 h in serum-supplemented or serum-free medium. As a control, cells were transfected with the respective vector alone without insert. The medium was collected and concentrated 10-fold. Cells were harvested after detachment with 10 mM EDTA for 10–15 min at 37 °C. A complete detachment of cells was confirmed by phase-contrast microscopy. Fifty microliters of 2× Laemmli sample buffer was added to the wells, and the ECM was scraped off. Samples of cell lysate, medium, and ECM were separately electrophoresed under reducing conditions (samples were boiled following the addition of 10% (v/v) 2-mercaptoethanol) on 12% SDS-polyacrylamide gels and transferred to enhanced chemiluminescence (ECL)-Hybond (Amersham Biosciences, Inc.). Western blotting was performed using either anti-FLAG M2 antibody (diluted 1:500, Sigma), anti-His (COOH-terminal) antibody (diluted 1:1000, Invitrogen) or antibody 4113 (diluted 1:300) depending on the construct used for transfection. Antibody binding was detected using the appropriate peroxidase-labeled second antibody followed by ECL using reagents from Amersham Biosciences, Inc.

For immunocytochemistry, COS-1 cells were grown on glass coverslips in 35-mm diameter wells (in 6-well plates) and transiently transfected as described above in serum-supplemented or serum-free medium. The medium was removed 48 h after transfections. The cells were washed three times on ice with cold PBS containing 1 mM $CaCl_2$ and 1 mM $MgCl_2$ and incubated for 1 h on ice with 1 ml of culture medium containing anti-FLAG M2 monoclonal antibody (diluted 1:300, Sigma) or anti-punctin rabbit antisera (diluted 1:100) with gentle shaking. Cells were washed four times for 3 min each with cold PBS, fixed in 4% paraformaldehyde (w/v in PBS) (Sigma) on ice for 30 min with gentle shaking and then washed three times with PBS at ambient temperature. To quench free aldehyde groups, cells were treated with 75 mM ammonium chloride, 20 mM glycine for 10 min at ambient temperature, washed with PBS, and then blocked with 0.05% Triton X-100, 2% normal goat serum in PBS (10 min at ambient temperature). Finally, sections were incubated with the species-appropriate Texas Red-labeled goat secondary antibody (Jackson ImmunoResearch Laboratories, West Grove, PA) prior to coverslip mounting in Vectashield containing 4′,6-diamidino-2-phenylindole (Vector Laboratories, Inc., Burlingame, CA). The following control-immunostaining experiments were performed. COS-1 cells transfected with the vector alone or untransfected COS-1 cells were stained with the above antibodies, or transfected cells were stained with preimmune serum from the rabbits in which the polyclonal antibodies were produced.

To co-stain punctin and the actin cytoskeleton, cells were stained with anti-FLAG or anti-punctin antibodies as described above with the exception that the secondary antibodies included incubation with Alexa 488-phalloidin at recommended dilutions (Molecular Probes). In double immunostaining experiments following the immunolocalization of FLAG or punctin as described above, cells were permeabilized with 0.1% Triton X-100 in PBS for 20 min prior to staining with (*a*) monoclonal antibody to vinculin (1:100 dilution, Sigma) in combination with antiserum 4113 for the detection of punctin or (*b*) polyclonal antibody to focal adhesion kinase (1:200 dilution, Upstate Biotechnology, Lake Placid, NY) in combination with anti-FLAG monoclonal antibody M2 (Sigma) for the detection of punctin. A Texas Red-labeled antibody (Jackson ImmunoResearch Laboratories) was used for the detection of punctin, and Alexa 488-conjugated antibody (Molecular Probes) was used for the detection of vinculin or focal adhesion kinase.

RESULTS

Cloning of Punctin cDNA—We identified a novel EST (GenBank™ accession number AA482392) derived from pooled human melanocyte, fetal heart, and pregnant uterus with homology to ADAMTS proteases. The 1.5-kb insert of the corresponding IMAGE clone 752797 contained a long ORF encoding an amino-terminal TS domain, a cysteine-rich domain, a cysteine-free spacer domain, and three tandem TS modules followed by a short acidic peptide and stop codon (Fig. 1*a*). The stop codon and 3′-untranslated sequence were independently confirmed by 3′-RACE (clone pSHTSL1s3, Fig. 1*a*) as well as by another EST (GenBank™ accession number W47029). The 3′-untranslated region encoded in IMAGE clone 752797 contained a consensus polyadenylation signal (AATTAAA) followed by a poly(A) tail 14 nucleotides downstream. Completion of the full-length coding sequences by 5′-RACE predicted a putative signal peptide upstream of the central TS domain. The signal peptide was preceded by a methionine codon within a satisfactory Kozak consensus sequence (A at −3, G at +4 relative to ATG) (20) although there was no upstream in-frame stop codon. The 5′ sequence obtained by RACE was subsequently validated by independently cloned human and mouse ESTs (Genbank™ accession numbers A1459225 for human EST and AK020115 for mouse EST). The continuity of the cDNA clones was confirmed by PCR amplification of the full-length punctin ORF from human skeletal muscle cDNA (see below) as well as by identification of the encoding exons arranged sequentially on human chromosome 9 (Celera Genomics, Rockville, MD).

Primary Structure of Punctin Predicts an ADAMTS-like Protein—The predicted full-length punctin protein contains 525 amino acids and has the typical domain structure of the ancillary noncatalytic regions of an ADAMTS protease (Fig. 1*a*). The mature secreted form of punctin is 497 amino acids with a molecular mass of 55,240 Da and a calculated pI of 6.2. Like the ADAMTS proteases, each domain in punctin has an even number of cysteine residues. This observation suggests that each domain may have internal disulfide bonds (17 such bonds are predicted in punctin), and that punctin consists of a series of independently-folded and disulfide-bonded domains. Punctin contains no other domains apart from those described previously in the ADAMTS family. The punctin sequence contains one motif for *N*-linked glycosylation (21) at Asn^{223} (-Asn-*X*-Ser/Thr-, where *X* is any amino acid except Pro) and also contains a total of 75 Thr and Ser residues, where *O*-linked glycosylation might occur. (Fig. 1*b*).

The overall punctin sequence is most similar to human AD-

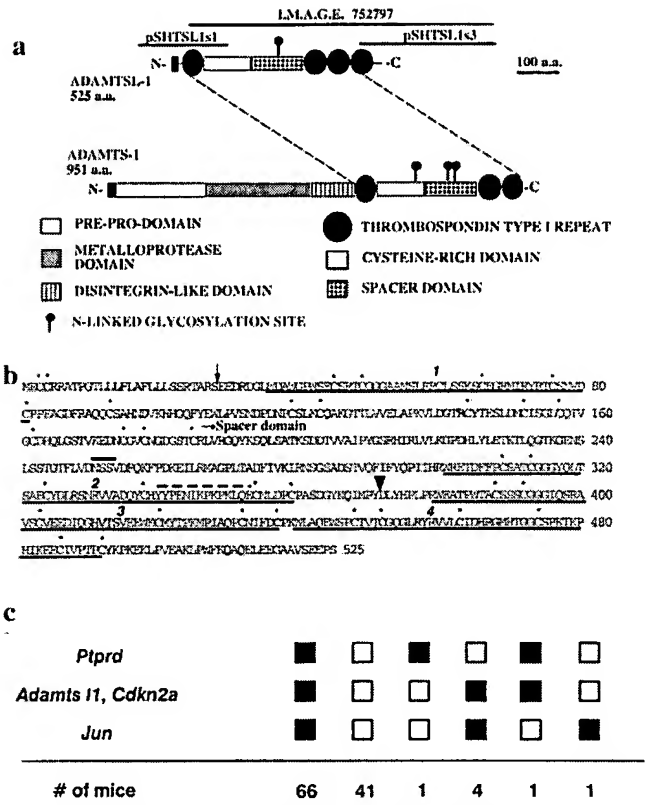


FIG. 1. *a*, domain organization of punctin/ADAMTSL-1 shown relative to ADAMTS-1, the prototypic ADAMTS. The cloning strategy used for determination of the complete primary structure is shown. The location of each cDNA clone relative to the protein domains indicates the regions it encodes. The key to the domains is shown at the *bottom* of the figure. *b*, the predicted amino acid sequence of punctin is shown using the single-letter amino acid code. TS modules are *underlined* with the *thick line* and are numbered sequentially from amino to carboxyl terminus. A consensus sequence for *N*-linked glycosylation is *overlined*. Cysteine residues are indicated by *asterisks*. The start of the spacer domain is indicated, the region between the NH_2-terminal TS domain and the spacer domain is the cysteine-rich domain. The *dashed line* indicates the peptide used for the generation of antibodies. The *arrow* indicates the signal peptidase cleavage site. The *arrowhead* indicates a putative proteolytic processing site between TS domains 2 and 3. *c*, segregation of ***Adamtsl1*** on mouse chromosome 4 in ((C3H/HeJ-*gld* × *M. spretus*) F_1 × C3H/HeJ-*gld*) interspecific backcross mice. *Filled boxes* represent the homozygous C3H pattern, and *open boxes* represent the F_1 pattern. The mapping of the reference loci in this interspecific cross has been previously described (15).

AMTSL-3 (68% identity, see below). Of the ADAMTS enzymes published to date, punctin is most similar to human ADAMTS-10 (35% identity). The punctin TS domains have a higher degree of similarity to other ADAMTS-like proteins and ADAMTS proteases than to TSP1 and TSP2. The greatest similarities, as indicated by percentage of identity of amino acid sequences identified by BLAST searches of the first TS domain of punctin to TS domains from various molecules, are as follows: human ADAMTSL-3, 80%; human ADAMTS-1, ADAMTS-6, and ADAMTS-10, 50%; mouse papilin, 47%; human ADAMTS-8, 44%; human ADAMTS-5, 42%; human TSP2, 40%; human TSP1, 38%. Like most TS domains in the ADAMTS family, punctin TS domains do not contain linear peptide sequences found in TSP1 that have been defined as heparin or CD-36 binding sequences, (22). They do not contain degenerate GAG binding sequences such as BBXB, where B is the basic amino acid and *X* is any amino acid (22).

Genomic Location of the Mouse and Human Punctin Genes and Tissue-specific Expression—The mapping of ***Adamtsl1*** in an interspecific cross resulted in the following most probable gene order (mean ± S.D.): ***Ptprd***-4.4 ± 2.0 centimorgan-***Adamtsl1, Cdkn2a***-1.8 ± 1.2 centimorgan-***Jun*** and placed ***Adamtsl1*** at a consensus position of 42.6 centimorgan on mouse chromosome 4 (Fig. 1c) in the vicinity of the interferon gene cluster. A search of the mouse genome data base (www.informatics.jax.org) did not reveal any pertinent genetic disorders near this locus.

The human-mouse homology maps (www3.ncbi.nlm.nih.gov/Omim/Homology/, accessed September 26, 2001) predict that the *ADAMTSL1* locus is on human chromosome 9p21-22. The predicted locus was confirmed by the analysis of the human genome sequence. The punctin ORF is encoded by 13 exons spanning >250 kb of genomic DNA mapping to 9p21.2-22.1. A search of the Online Mendelian Inheritance in Man site (www3.ncbi.nlm.nih.gov/Omim/) revealed three unsolved human disorders in the vicinity of the *ADAMTSL1* locus. Diaphyseal medullary stenosis with malignant fibrous histiocytoma (MIM112250) is linked to 9p22-p21, Friedreich's ataxia 2 (MIM601992) is linked to 9p23-p11, and neuropathy, distal hereditary motor, Jerash type (MIM605726) are linked to 9p21.1-p12.

ADAMTSL1 is primarily expressed in human and mouse skeletal muscle with a major message size of ~7.0 kb in both species (Fig. 2). A minor messenger RNA species of ~1.0 kb was also seen in some human tissues (Fig. 2, skeletal muscle, heart, colon, kidney, and liver). Expression was not detected in brain, colon, thymus, spleen, placenta, small intestine, lung, testis, ovary, or peripheral blood leukocytes.

Expression and Characterization of Recombinant Punctin—Punctin expressed in High-Five™ cells with tandem COOH-terminal V5 and 6× His epitopes was secreted into the conditioned medium of adherent as well as suspension cultures. Punctin was detected by antibody 4113 and anti-epitope tag antibodies as a ~60-kDa band under reducing conditions. It was substantially purified from the culture medium using Ni-chromatography (Fig. 3*a*). The purification scheme yielded a maximum of 200 μg/liter purified protein as determined by amino acid analysis. Electrophoresis and Western blotting of concentrated punctin preparations frequently demonstrated additional bands of molecular mass (~120 and ~180 kDa, data not shown), suggesting the formation of dimers and trimers at high concentrations.

The conformation of punctin appears to be maintained by disulfide bonds as evidenced by more rapid migration in SDS-PAGE under nonreducing conditions than under reducing conditions (Fig. 3*b*). Furthermore, on Western blots under nonreducing conditions, the protein was not detectable with antibody 4113 (data not shown), suggesting that the peptide epitope was not accessible without reduction of disulfide bonds. A mass analysis of His-tagged punctin by MALDI-TOF MS yielded a broad peak suggesting that the 60-kDa gel band contained major molecular species of 61,935 ± 595 and 60,873 ± 295 Da, respectively. LC-ESMS analyses of the intact protein defined more precisely the major molecular species to be 61,970 ± 6 and 62,131 ± 5, which are, respectively, 1979 and 2140 Da larger than the calculated mass (59,991) of tagged punctin based on amino acid sequence. NH_2-terminal sequencing of the polyvinylidene difluoride-immobilized 60-kDa protein revealed a single sequence, which commenced at Glu^{29} (*i.e.* Glu-Glu-Asp-Arg-Asp-Gly and so on).

Recombinant Punctin Is Glycosylated—Two closely spaced punctin bands were resolved by Western blot analysis of conditioned medium or purified protein, although Coomassie Blue staining of purified punctin always demonstrated a single band (Fig. 3*a*). A periodic acid-Schiff-based method of staining carbohydrate chains suggested that recombinant punctin is a gly-

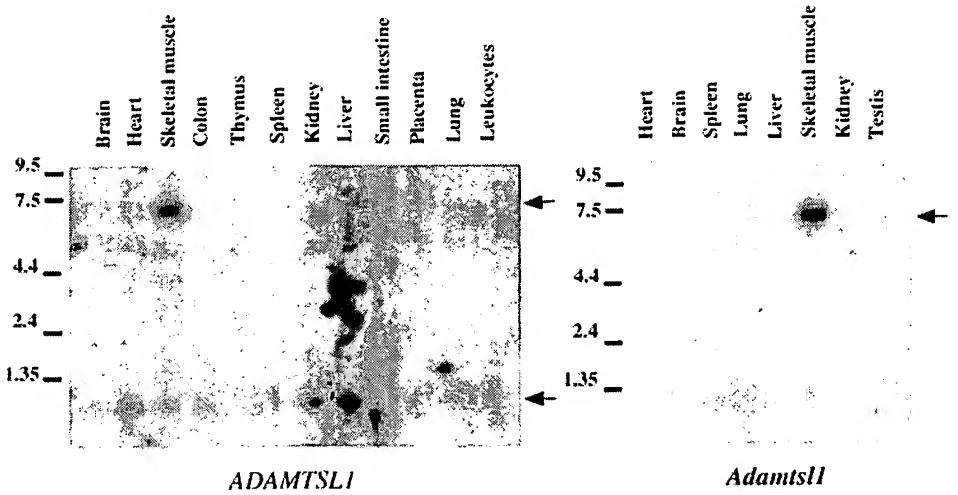


FIG. 2. **Northern analysis of expression of *ADAMTSL1* (*left*) and *Adamtsl1* (*right*) in adult human and mouse tissues, respectively.** Kilobase markers of RNA are shown at the *left* of each autoradiogram, and tissue origin is indicated above each lane. Hybridizing transcripts are indicated by *arrows*.

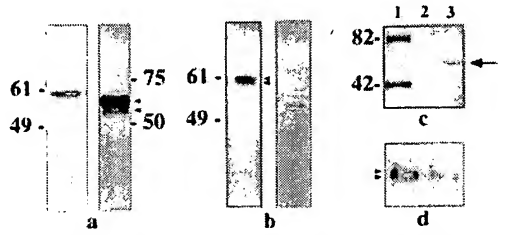


FIG. 3. **Analysis of epitope-tagged punctin purified by Ni-chromatography from insect cell culture medium.** *a*, Coomassie Blue (Simply Blue Safe Stain) staining of purified recombinant punctin on reducing SDS-PAGE (*left lane*) and Western blot analysis with anti-punctin antibody 4113 (*right lane*). *b*, Western blot analysis using anti-His tag monoclonal antibody on reducing (*left lane*) and nonreducing SDS-PAGE (*right lane*). *c*, glycoprotein staining of recombinant punctin (*lane 2* contains 0.6 μg, and *lane 3* contains 3 μg) using the periodic acid-Schiff procedure. Glycosylated CandyCane™ markers (1 μg/band) stained similarly are in *lane 1*. The *arrow* indicates stained punctin. *d*, Western analysis of culture medium from insect cell cultures treated without (*left lane*) or with (*right lane*) tunicamycin A for 72 h. Each lane contains 2.8 μg of total protein. *Double arrowheads* are used to indicate two molecular species seen on Western blots.

coprotein (Fig. 3*c*), and mass spectrometry demonstrated multiple molecular species consistent with variable glycosylation. Treatment of recombinant protein with peptide *N*-glycosidase F did not result in a perceptible decrease in molecular mass, although the intensity of glycoprotein staining was decreased (data not shown). Culture medium from tunicamycin-treated cells exhibited only a single punctin species as demonstrated by Western blotting (Fig. 3*d*). The difference (161 Da) between the LC-ESMS-observed masses of the major punctin molecular species (61,970 and 62,131 Da) is close to the in-chain chemical average mass of a oligosaccharide residue (Hex, 162). Minor molecular species were also apparent by LC-ESMS analysis, which differed by mass increments that approximated the in-chain chemical average mass of oligosaccharide residues (*e.g.* Hex, 162; HexNAc, 203; NeuAc, 291). For a further analysis, tryptic digests of the protein were examined by analytical LC-ESMS using stepped collision energy scanning to produce carbohydrate-specific marker ions. Glycopeptides were detected including molecular species with masses of 5881.4 ± 0.4 and 6171.2 ± 0.2 Da. The mass difference (289.8 Da) between these observed glycopeptides appears to correspond to the in-chain chemical average mass of *N*-acetylneuraminic acid (NeuAc, 291). Taken together, these data indicated that punctin is glycosylated, although specific glycopeptides have yet to be characterized fully. Approximately 65% of the amino acid sequence in punctin was identified by peptide mass mapping including the NH_2-terminal tryptic peptide (Glu^{29}-Arg^{47}), verifying that the target protein has been expressed. Based on the difference between the observed and calculated masses of intact punctin, the recombinant protein contains approximately 3–4% carbohydrate by weight.

During purification of punctin in the absence of protease inhibitors, additional components of ~40 and 20 kDa, respectively, were detected on Coomassie Blue-stained gels (data not shown). The 40-kDa band contained two molecular species with measured masses of 38,409 ± 115 and 39,456 ± 156 Da, respectively, as determined by MALDI-TOF MS. The NH_2-terminal sequencing of these bands yielded the same amino terminus as the full-length punctin. The ~20-kDa fragment exhibited an NH_2-terminal sequence ^{372}DLYHPL, indicating that the fragment is from the carboxyl terminus. The addition of 1 mM phenylmethylsulfonyl fluoride to culture medium effectively prevented this proteolysis, suggesting that it was effected by a serine protease.

Visualization of Punctin by Rotary Shadowing—Rotary shadowing of purified recombinant punctin demonstrated a hatchet-shaped or comma-shaped molecule 30–40 μm in length (Fig. 4). Punctin consists of a single globular domain of 10–20 μm in size with a short linear segment at one end. Most of the visualized protein was in monomeric form (Fig. 4). Occasional aggregates with the appearance of dimers and trimers were seen but have not yet been resolved in detail.

Expression and Localization of Punctin in Transfected COS-1 Cells—Transfected cells were stained without fixation or permeabilization and on ice (live staining) to prevent the detection of intracellular punctin or endocytosed antibody, respectively. Under these conditions, punctin was localized underneath the cells (*i.e.* adjacent to their ventral surface) in the substratum laid down on plastic. The staining pattern was punctate (Fig. 5, *a–d*) and was preferentially located toward the periphery of the cells (Fig. 5, *a*, *b*, and *d*) and under cellular processes (Fig. 5*c*). The punctin deposits were of submicron dimension, although fluorescent signals from closely located deposits were frequently merged suggesting larger aggregates. Transfected cells had minimal or no staining on the dorsal cell surface. Punctin was not seen in the substratum in areas not corresponding to the cells. If cells were detached with 10 mM EDTA prior to staining, "footprints" of transfected cells were retained on the substratum with a similar staining pattern as under intact

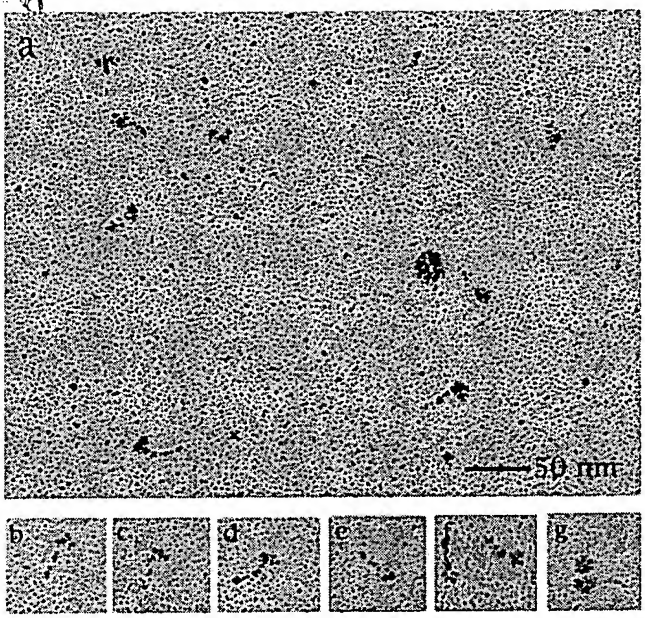


FIG. 4. **Rotary shadowing of recombinant punctin.** *a*, overview. *b–g*, images of individual punctin molecules. *Scale bar* in *panel a* indicates molecular dimensions in all panels.

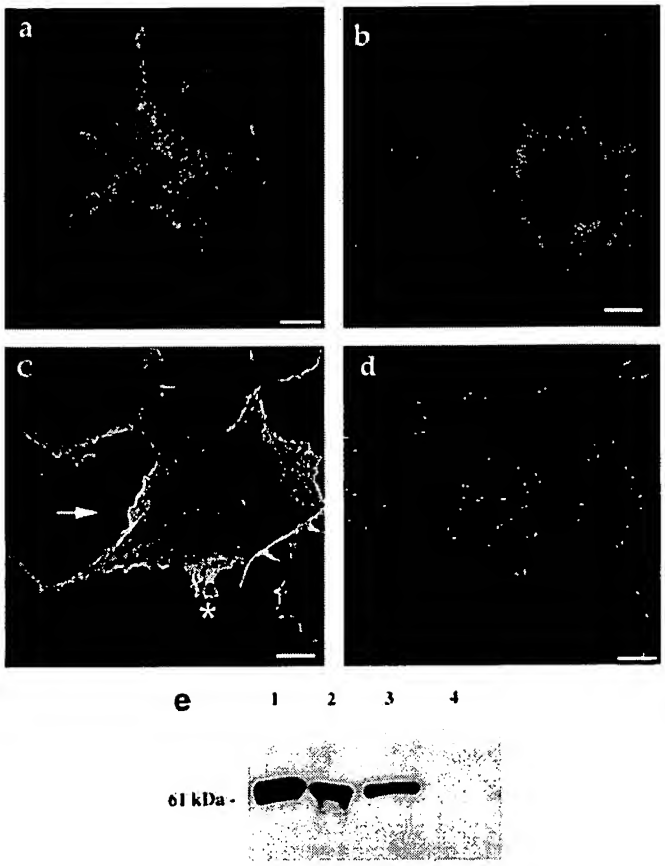


FIG. 5. *a–d*, confocal laser-scanning microscopy of COS-1 cells following transient transfection with *ADAMTSL1* expression constructs and immunocytochemistry. Untransfected cells are visible in *a* and *b*. Scale bar (10 μm) is shown at *lower right* of each panel. *a* and *b*, punctate staining of FLAG-tagged punctin (*red*) in nonpermeabilized cells visualized with anti-FLAG M2 antibody. Nuclei are *blue* 4′,6-diamidino-2-phenylindole. *c*, relationship of punctin staining (*red*) visualized with anti-FLAG M2 monoclonal antibody to cellular actin as visualized by phalloidin staining (*green*). The *asterisk* indicates a cellular protrusion that does not have underlying punctin, and the *arrow* indicates punctin immunolocalization without an overlying cellular process. *d*, relationship of punctin (*red*) visualized with anti-punctin antiserum 4113 to vinculin staining (*green*) as shown by confocal imaging and overlay of single-color images from a double-stained cell. *e*, Western blot analysis of cell lysates (*lane 1*), medium (*lane 2*), and ECM (*lane 3*) from transfected COS-1 cells using an anti-His tag monoclonal antibody. Cell lysates from untransfected COS-1 cells are shown in *lane 4*. Molecular mass is indicated on the *left*.

cells. Staining was seen in some areas not covered with cell processes. In other areas, there were cell processes without underlying punctin (Fig. 5*c*). We interpret this finding to result from cellular motility (*i.e.* withdrawal of existing processes and the formation of new ones). Identical results were obtained with anti-FLAG monoclonal antibody or antibody 4113. Fig. 5, *a–c*, shows staining of FLAG-tagged protein using the FLAG M2 monoclonal antibody, and Fig. 5*d* shows staining with anti-punctin antiserum 4113. Similar staining patterns were seen whether cells were grown in the presence or absence of serum and using tagged or untagged proteins (data not shown).

Double staining for vinculin (Fig. 5*d*) or focal adhesion kinase (data not shown), components of focal contacts, indicated that punctin staining did not correspond to sites of focal contacts. No staining was visible in control experiments, *i.e.* in untransfected COS cells, cells transfected with vector alone, cells stained without a primary antibody, or cells stained with preimmune serum as control.

On Western blots, we found reactive protein bands of the expected size (58–60 kDa for untagged punctin and 62–64 kDa for the His-tagged or FLAG-tagged forms) in the medium, cell layer, and the underlying substratum or ECM of transfected COS-1 cells (Fig. 5*e*). In contrast, cells transfected with vector alone (Fig. 5*e*) or untransfected cells (data not shown) did not show a reactive band. As controls, preimmune serum from the rabbits in which anti-Punctin antibodies were generated did not produce immunoreactivity on Western blots (data not shown).

DISCUSSION

Punctin/ADAMTSL-1 Is a Novel ADAMTS-like Secreted Protein Belonging to a Distinct ADAMTSL Family of Proteins—In addition to missing the catalytic domain, the ADAMTS-like proteins (see below) do not possess disintegrin-like domains. This finding suggests that the disintegrin-like domain and catalytic domain may represent a functionally coupled protease domain in ADAMTS enzymes. Further evidence for this comes from the identification of other proteins with a predicted structure similar to punctin. Following the complete cloning of punctin/ADAMTSL-1, we became aware of a second such molecule encoded by the *KIAA0605* gene (GenBank™ accession number AB011177) that we designated as ADAMTSL-2 (23). We have cloned a third ADAMTS-like protein, ADAMTSL-3 (GenBank™ accession number AF237652).[3] Therefore, punctin belongs to a distinct protein family. ADAMTSL-2 and ADAMTSL-3 differ from punctin in their greater length (951 and 1690 amino acids, respectively) and also have more TS domains (6 and 10, respectively). These molecules will be described in greater detail in subsequent publications. In contrast to ADAMTSL-2 and ADAMTSL-3, which are quite widely expressed,[4] punctin/ADAMTSL-1 is selectively expressed in muscle.

Other secreted ECM molecules such as lacunin and papilin also contain the ancillary domains of the ADAMTS family in the precise order as punctin. However, punctin is more closely related to ADAMTSL-3 and some ADAMTS proteases than it is to mouse papilin (32% identity). Lacunin is a basement membrane glycoprotein in the moth *Manduca sexta* (24). Lacunin has the structure of ADAMTSL including seven TS modules as

[3] N. Moore, B. Anand-Apte, and S. Apte, unpublished data.
[4] S. Apte, unpublished data.

well as a single COOH-terminal protease and lacunin domain. In addition, it contains 13 repeats of a novel lagrin domain, 11 Kunitz inhibitor domains, 2 antistasin-like domains, 1 serine protease inhibitor domain, and 2 immunoglobulin domains. Lacunin localizes to the basal lamina of the moth wing (24). Papilin from *Drosophila melanogaster* may be an ortholog of *M. sexta* lacunin, because the two molecules are similar in their domain content, organization, and primary sequence. Papilin is also a basement membrane protein (25). Although these invertebrate proteins have numerous protease inhibitor domains, mammalian papilin contains substantially fewer such domains (25).

Characterization of Recombinant Punctin from Insect Cells—Our experimental data support the likelihood that recombinant punctin is disulfide-bonded. First, its electrophoretic mobility is greater under nonreducing conditions. Second, the punctin epitope is masked under nonreducing conditions. Third, rotary shadowing demonstrated a molecule with a specific and consistent conformation. Limited proteolysis within the linker peptide, connecting TS domains 2 and 3 assigned to the Tyr^{371}-Asp^{372} peptide bond (Fig. 1*b*) by a putative serine protease, indicates that there may be a proteolytically susceptible exposed region between the two disulfide-bonded TS domains. It is not yet known whether this is a physiologically relevant processing or whether it is an artifact that is unique to this expression system. The processing event releases the two COOH-terminal TS domains of punctin. Because proteolytically derived fragments of many secreted proteins have distinctive functions, it will be interesting to investigate whether specific functions are associated with the ~40- and ~20-kDa fragments.

A mass measurement of epitope-tagged recombinant punctin by MALDI-TOF MS and LC-ESMS revealed that purified punctin contained multiple species of higher than the predicted mass. Edman degradation indicated that all these species had the same amino terminus. Further MS analysis, glycoprotein staining, and culture in the presence of tunicamycin A confirm that punctin contains *N*-linked sugars but do not exclude the presence of *O*-linked sugar. Significant alteration of mobility was not seen after peptide *N*-glycosidase F treatment, suggesting that the *N*-linked carbohydrate may be resistant to complete enzymatic removal (26).

Rotary shadowing is useful for demonstrating the physical conformation of a molecule as well as the existence of oligomeric complexes (27–29). The data we have obtained for punctin are relevant to the ADAMTS, lacunin, and papilin. They can be extrapolated to represent the structure of the ancillary domains of an ADAMTS enzyme and the "papilin cassette" (25) and provide the first insight into the conformation of these domain assemblies. Many ECM proteins exist as oligomers. This observation may also be the case with punctin, because rotary shadowing electron microscopy and gel electrophoresis occasionally suggested the presence of dimers and trimers. We anticipate that rotary shadowing will be useful for future studies to investigate punctin oligomerization and interactions of punctin with putative ECM ligands.

Punctin Is an ECM Glycoprotein That Binds to the Cell Substratum in a Spatially Specific Manner—Nontransformed cells in culture require a substratum for attachment, spreading, and migration. The substratum present on an unmodified plastic tissue culture surface is derived from the cells themselves as well as from proteins in serum-supplemented culture medium (30–32). Quantitatively significant components of the cell substratum are laminin, fibronectin, vitronectin, collagen, tenascin, PG-M or versican (a chondroitin sulfate proteoglycan), perlecan (a heparan sulfate proteoglycan), hyaluronan, and tissue inhibitor of metalloproteases-3 (30–37). Punctin shares the subcellular distribution of molecules that do not generally co-localize with focal contacts (*e.g.* versican, hyaluronan, and tenascin) (31, 37). Because punctin is left behind in the ECM after cell detachment with EDTA, we conclude that when expressed in COS-1 cells, punctin binds a component of the ECM. Punctin in culture medium may reflect an excess of more than that which can bind to the substratum or indicate secretion from the free surface of the cell. Punctin does not bind to ECM between the cells, indicating that the punctin ligand is absent from these regions. Because similar staining was seen under serum-supplemented as well as under serum-free culture conditions, it is probable that the ECM binding partner of punctin is a molecule produced by COS-1 cells but not one derived from fetal bovine serum.

Significance of Punctin and the ADAMTS-like Family—Molecules comprising ancillary domains of metalloproteases may be generated in biological systems by proteolytic processing or through alternative splicing of protease genes. Brooks *et al.* (38) found that the proteolytically generated hemopexin domain of MMP-2 circulated in serum and bound to the integrin $\alpha_v\beta_3$. This MMP-2 fragment inhibited angiogenesis by preventing membrane targeting of MMP-2 (38). So far, there are no known examples of ADAMTS-like proteins generated as splice variants of *ADAMTS* genes. The discovery of punctin demonstrates for the first time the existence of molecules closely resembling the ancillary domains of ADAMTS that are generated as distinct gene products.

The resemblance of ADAMTSL to ADAMTS suggests a functional relationship between these two groups of molecules. From studies on ADAMTS-1 (39) and ADAMTS-2 (40), it is known that the ancillary domains are required to bind and cleave substrates. ADAMTSL may offer a potential mechanism of ADAMTS regulation via one of several possible mechanisms. As a result of noncompetitive inhibition of ADAMTS-2, an inhibitory role has been shown for *Drosophila* papilin (25). Another possibility is that punctin may compete with ADAMTS for its substrates and protect the substrates from cleavage. The isolated MMP-2 hemopexin domain represents one such example. In a second example, a truncated nonenzymatic version of ADAM-17 was shown to have a dominant negative effect on the activation of tumor necrosis factor-α (41). An intriguing possibility is that the ADAMTS-like proteins may be enhancers of the ADAMTS proteases. For example, the procollagen C-proteinase enhancer protein (42) contains two domains homologous to those found in the C-proteinase that are instrumental in binding to the carboxyl propeptide of procollagen I and enhancing its removal (43). Very little is currently known about the regulation of ADAMTS proteases following their activation, and it is possible that the ADAMTS-like proteins may provide a novel general principle of regulation.

Acknowledgments—We thank C. Kassuba for editing the manuscript, Vincent C. Hascall, Tom Tallant, Seng Hui Low, and Thomas Weimbs for helpful discussion, members of the Apte laboratory for critical reading of the manuscript, Satya Yadav for protein sequencing, Monique Ross for assistance with cloning, Karen West for amino acid analysis, and Judy Drazba for help with confocal imaging.

REFERENCES

1. Murphy, G., and Knauper, V. (1997) *Matrix Biol.* **15**, 511–518
2. Blobel, C. P. (1997) *Cell* **90**, 589–592
3. Kuno, K., Kanada, N., Nakashima, E., Fujiki, F., Ichimura, F., and Matsushima, K. (1997) *J. Biol. Chem.* **272**, 556–562
4. Hurskainen, T. L., Hirohata, S., Seldin, M. F., and Apte, S. S. (1999) *J. Biol. Chem.* **274**, 25555–25563
5. Wallon, U. M., and Overall, C. M. (1997) *J. Biol. Chem.* **272**, 7473–7481
6. Steffensen, B., Bigg, H. F., and Overall, C. M. (1998) *J. Biol. Chem.* **273**, 20622–20628
7. Bein, K., and Simons, M. (2000) *J. Biol. Chem.* **275**, 32167–32173
8. Bornstein, P., Armstrong, L. C., Hankenson, K. D., Kyriakides, T. R., and Yang, Z. (2000) *Matrix Biol.* **19**, 557–568

9. Overall, C. M., King, A. E., Sam, D. K., Ong, A. D., Lau, T. T., Wallon, U. M., De Clerck, Y. A., and Atherstone, J. (1999) *J. Biol. Chem.* **274,** 4421–4429
10. McQuibban, G. A., Gong, J. H., Tam, E. M., McCulloch, C. A., Clark-Lewis, I., and Overall, C. M. (2000) *Science* **289,** 1202–1206
11. Zhang, X. P., Kamata, T., Yokoyama, K., Puzon-McLaughlin, W., and Takada, Y. (1998) *J. Biol. Chem.* **273,** 7345–7350
12. Zolkiewska, A. (1999) *Exp. Cell Res.* **252,** 423–431
13. Kuno, K., and Matsushima, K. (1998) *J. Biol. Chem.* **273,** 13912–13917
14. Georgiadis, K. E., Hirohata, S., Seldin, M. F., and Apte, S. S. (1999) *Genomics* **62,** 312–315
15. Watson, M. L., and Seldin, M. F. (1994) *Methods Mol. Genet.* **5,** 369–387
16. Crabb, J. W., West, K. A., Dodson, W. S., and Hulmes, J. D. (1997) in *Current Protocols in Protein Science* (Coligan, J. E., Ploegh, H. L., Smith, J. A., and Speicher, D. W., eds) pp. 11.9.1–11.9.42, John Wiley & Sons, Inc., New York
17. Kapron, J. T., Hilliard, G. M., Lakins, J. N., Tenniswood, M. P., West, K. A., Carr, S. A., and Crabb, J. W. (1997) *Protein Sci.* **6,** 2120–2133
18. Crabb, J. W., Nie, Z., Chen, Y., Hulmes, J. D., West, K. A., Kapron, J. T., Ruuska, S. E., Noy, N., and Saari, J. C. (1998) *J. Biol. Chem.* **273,** 20712–20720
19. Sakai, L. Y., and Keene, D. R. (1994) *Methods Enzymol.* **245,** 29–52
20. Kozak, M. (1989) *J. Cell Biol.* **108,** 229–241
21. Gavel, Y., and von Heijne, G. (1990) *Protein Eng.* **3,** 433–442
22. Bornstein, P., and Sage, E. H. (1994) *Methods Enzymol.* **245,** 62–85
23. Nagase, T., Ishikawa, K., Miyajima, N., Tanaka, A., Kotani, H., Nomura, N., and Ohara, O. (1998) *DNA Res.* **5,** 31–39
24. Nardi, J. B., Martos, R., Walden, K. K., Lampe, D. J., and Robertson, H. M. (1999) *Insect Biochem. Mol. Biol.* **29,** 883–897
25. Kramerova, I. A., Kawaguchi, N., Fessler, L. I., Nelson, R. E., Chen, Y., Kramerov, A. A., Kusche-Gullberg, M., Kramer, J. M., Ackley, B. D., Sieron, A. L., Prockop, D. J., and Fessler, J. H. (2000) *Development* **127,** 5475–5485
26. Tretter, V., Altmann, F., and Marz, L. (1991) *Eur. J. Biochem.* **199,** 647–652
27. Holden, P., Meadows, R. S., Chapman, K. L., Grant, M. E., Kadler, K. E., and Briggs, M. D. (2001) *J. Biol. Chem.* **276,** 6046–6055
28. Maddox, B. K., Sakai, L. Y., Keene, D. R., and Glanville, R. W. (1989) *J. Biol. Chem.* **264,** 21381–21385
29. Olin, A. I., Morgelin, M., Sasaki, T., Timpl, R., Heinegard, D., and Aspberg, A. (2001) *J. Biol. Chem.* **276,** 1253–1261
30. Hayman, E. G., Oldberg, A., Martin, G. R., and Ruoslahti, E. (1982) *J. Cell Biol.* **94,** 28–35
31. Hayman, E. G., Pierschbacher, M. D., Suzuki, S., and Ruoslahti, E. (1985) *Exp. Cell Res.* **160,** 245–258
32. Ohji, M., Sundar Raj, N., Hassell, J. R., and Thoft, R. A. (1994) *Invest. Ophthalmol. Visual Sci.* **35,** 479–485
33. Snow, A. D., Sekiguchi, R., Nochlin, D., Fraser, P., Kimata, K., Mizutani, A., Arai, M., Schreier, W. A., and Morgan, D. G. (1994) *Neuron* **12,** 219–234
34. Singer, I. I., II, Scott, S., Kawka, D. W., and Hassell, J. R. (1987) *Exp. Cell Res.* **173,** 558–571
35. Isogai, Z., Shinomura, T., Yamakawa, N., Takeuchi, J., Tsuji, T., Heinegard, D., and Kimata, K. (1996) *Cancer Res.* **56,** 3902–3908
36. Staskus, P. W., Masiarz, F. R., Pallanck, L. J., and Hawkes, S. P. (1991) *J. Biol. Chem.* **266,** 449–454
37. Yamagata, M., Saga, S., Kato, M., Bernfield, M., and Kimata, K. (1993) *J. Cell Sci.* **106,** 55–65
38. Brooks, P. C., Silletti, S., von Schalscha, T. L., Friedlander, M., and Cheresh, D. A. (1998) *Cell* **92,** 391–400
39. Kuno, K., Terashima, Y., and Matsushima, K. (1999) *J. Biol. Chem.* **274,** 18821–18826
40. Colige, A., Sieron, A. L., Li, S. W., Schwarze, U., Petty, E., Wertelecki, W., Wilcox, W., Krakow, D., Cohn, D. H., Reardon, W., Byers, P. H., Lapiere, C. M., Prockop, D. J., and Nusgens, B. V. (1999) *Am. J. Hum. Genet.* **65,** 308–317
41. Solomon, K. A., Pesti, N., Wu, G., and Newton, R. C. (1999) *J. Immunol.* **163,** 4105–4108
42. Takahara, K., Kessler, E., Biniaminov, L., Brusel, M., Eddy, R. L., Jani-Sait, S., Shows, T. B., and Greenspan, D. S. (1994) *J. Biol. Chem.* **269,** 26280–26285
43. Hulmes, D. J., Mould, A. P., and Kessler, E. (1997) *Matrix Biol.* **16,** 41–45